

German Data Forum (RAtSWD) [Hrsg.]

**Building on progress. Expanding the research infrastructure for the social, economic, and behavioral sciences. Vol. 2**

Opladen ; Farmington Hills, Mich. : Budrich UniPress 2010, 627 - 1238 S.



Quellenangabe/ Reference:

German Data Forum (RAtSWD) [Hrsg.]: Building on progress. Expanding the research infrastructure for the social, economic, and behavioral sciences. Vol. 2. Opladen ; Farmington Hills, Mich. : Budrich UniPress 2010, 627 - 1238 S. - URN: urn:nbn:de:0111-opus-35747 - DOI: 10.25656/01:3574

<https://nbn-resolving.org/urn:nbn:de:0111-opus-35747>

<https://doi.org/10.25656/01:3574>

in Kooperation mit / in cooperation with:



**Nutzungsbedingungen**

Dieses Dokument steht unter folgender Creative Commons-Lizenz:  
<http://creativecommons.org/licenses/by-nc-nd/3.0/de/deed> - Sie dürfen das Werk bzw. den Inhalt unter folgenden Bedingungen vervielfältigen, verbreiten und öffentlich zugänglich machen: Sie müssen den Namen des Autors/Rechteinhabers in der von ihm festgelegten Weise nennen. Dieses Werk bzw. dieser Inhalt darf nicht für kommerzielle Zwecke verwendet werden und es darf nicht bearbeitet, abgewandelt oder in anderer Weise verändert werden.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

**Terms of use**

This document is published under following Creative Commons-License:  
<http://creativecommons.org/licenses/by-nc-nd/3.0/de/deed.en> - You may copy, distribute and transmit, adapt or exhibit the work in the public as long as you attribute the work in the manner specified by the author or licensor. You are not allowed to make commercial use of the work or its contents. You are not allowed to alter, transform, or change this work in any other way.

By using this particular document, you accept the above-stated conditions of use.



**Kontakt / Contact:**

peDOCS  
DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation  
Informationszentrum (IZ) Bildung  
E-Mail: [pedocs@dipf.de](mailto:pedocs@dipf.de)  
Internet: [www.pedocs.de](http://www.pedocs.de)

RatSWD.

German Data Forum (ed.)

# Building on Progress

Expanding the Research Infrastructure  
for the Social, Economic, and  
Behavioral Sciences

vol | ume | 2

Building on Progress  
Expanding the Research Infrastructure for the  
Social, Economic, and Behavioral Sciences



# Building on Progress

Expanding the Research Infrastructure  
for the Social, Economic, and Behavioral  
Sciences

edited by the  
German Data Forum (RatSWD)

Vol. 2

Budrich UniPress Ltd.  
Opladen & Farmington Hills, MI 2010

You must not circulate this book in any other binding or cover and you must impose this same condition on any acquirer.

A CIP catalogue record for this book is available from  
Die Deutsche Bibliothek (The German Library)

© 2010 by Budrich UniPress Ltd. Opladen & Farmington Hills  
[www.budrich-unipress.eu](http://www.budrich-unipress.eu)

**ISBN 978-3-940755-58-2**  
**2 Vols. in slipcase**

© These volumes are published by Budrich UniPress and are published under the Creative Commons licence:  
[http://creativecommons.org/licenses/by-nc-nd/3.0/de/deed.en\\_GB](http://creativecommons.org/licenses/by-nc-nd/3.0/de/deed.en_GB)  
You are free to copy, distribute, display, and perform the work as long as you tribute it to the original author/editors. No commercial use, or changes unless approved by Budrich UniPress.



Both volumes may be downloaded free of charge at the Budrich UniPress website  
[www.budrich-unipress.com](http://www.budrich-unipress.com)  
The two volumes in slipcase may be ordered from your local bookseller or directly from the Budrich UniPress.

Budrich UniPress Ltd.  
Stauffenbergstr. 7. D-51379 Leverkusen Opladen, Germany  
28347 Ridgebrook. Farmington Hills, MI 48334. USA  
[www.budrich-unipress.eu](http://www.budrich-unipress.eu)

Jacket illustration by Walburga Fichtner, Cologne, Germany  
Printed in Europe on acid-free paper by  
paper & tinta, Warsaw, Poland

## Table of Contents:

### Volume 1:

<b>PREFACE BY THE FEDERAL MINISTRY OF EDUCATION AND RESEARCH (BMBF)</b> .....	i-iii
<b>EDITORS</b> .....	9
<b>PREFACE BY THE GERMAN DATA FORUM (RATSWD)</b> .....	11
<b>RECOMMENDATIONS</b> .....	17
<b>EXECUTIVE SUMMARIES</b> .....	41
<b>ADVISORY REPORTS</b> .....	95
TOWARDS AN IMPROVED RESEARCH INFRASTRUCTURE FOR THE SOCIAL SCIENCES: FUTURE DEMANDS AND NEEDS FOR ACTION .....	97
1. Providing a Permanent Institutional Guarantee for the German Information Infrastructure... <i>Johann Hahlen</i> .....	97
2. The European Dimension... <i>Klaus Reeh</i> .....	115
3. The Role of the German Research Foundation... <i>Eckard            Kämper and Manfred Nießen</i> .....	127
4. Providing Data on the European Level... <i>Peter Elias</i> .....	139
5. Infrastructure for High-Quality and Large-Scale Surveys. Cooperation between Academic Research and Private-Sector Agencies... <i>Peter Ph. Mohler and Bernhard von Rosenbladt</i> .....	155
6. The Availability of Market Research Data and its Potential for Use in Empirical Social and Economic Research... <i>Erich            Wiegand</i> .....	175
<b>PROGRESS SINCE 2001 AND CURRENT STATE</b> .....	189
1. The Recommendations of the 2001 KVI Report and Their Realizations up to 2008... <i>Gabriele Rolf-Engel</i> .....	189
2. Access to Microdata from Official Statistics... <i>Stefan Bender,            Ralf Himmelreicher, Sylvia Zühlke and Markus Zwick</i> .....	215
3. Publicly Financed Survey Data: Access and Documentation <i>Wolfgang Jagodzinski and Christof Wolf</i> .....	231

4. Teaching and Statistical Training... <i>Ulrich Rendtel</i> .....	251
5. e- Infrastructure for the Social Sciences... <i>Ekkehard Mochmann</i> .....	265
CURRENT STATE OF AND FUTURE DEMANDS IN DIFFERENT FIELDS.....	287
<b>I. (NEW) DATA TYPES AND THEIR POTENTIALS</b> .....	287
1. Macroeconomic Data... <i>Ulrich Heilemann</i> .....	289
2. Interdisciplinary Longitudinal Surveys... <i>Jürgen Schupp and Joachim Frick</i> .....	303
3. Geodata... <i>Tobia Lakes</i> .....	325
4. Regional Data... <i>Gerd Grözinger and Wenzel Matiaske</i> .....	341
5. Genetically Sensitive Sample Designs... <i>Frank M. Spinath</i> .....	353
6. Biological Variables in Social Surveys... <i>Rainer Schnell</i> .....	367
7. Administrative Transaction Data... <i>Julia Lane</i> .....	383
8. Transaction Data: Commercial Transaction Surveys and Test Market Data... <i>Bernhard Engel</i> .....	401
9. Time Use and Time Budgets... <i>Joachim Merz</i> .....	413
<b>II. METHODS</b> .....	451
1. Survey Methodology: International Developments... <i>Frauke Kreuter</i> .....	453
2. Improving Research Governance through Use of the Total Survey Error Framework... <i>Marek Fuchs</i> .....	471
3. Metadata... <i>Arofan Gregory, Pascal Heus and Jostein Ryssevik</i> .....	487
4. Paradata... <i>Frauke Kreuter and Carolina Casas-Cordero</i> .....	509
5. Record Linkage from a Technical Point of View... <i>Rainer Schnell</i> .....	531
6. Experiments, Surveys, and the Use of Representative Samples as Reference Data... <i>Thomas Siedler and Bettina Sonnenberg</i> .....	547
7. Experimental Economics... <i>Simon Gächter</i> .....	563
8. Experience Sampling... <i>Michaela Riediger</i> .....	581
9. Virtual Worlds as Petri Dishes for the Social and Behavioral Sciences... <i>Edward Castronova and Matthew Falk</i> .....	595



10. Qualitative Interviewing of Respondents in Large Representative Surveys... <i>Olaf Groh-Samberg and Ingrid Tucci</i> .....	607
---	-----

## Volume 2:

<b>III. DATA PROTECTION AND RESEARCH ETHICS</b> .....	627
1. Data Protection and Statistics – A Dynamic and Tension- Filled Relationship... <i>Peter Schaar</i> .....	629
2. Record Linkage from the Perspective of Data Protection... <i>Rainer Metschke</i> .....	643
3. New Methods in Human Subjects Research: Do We Need a New Ethics?... <i>Karsten Weber</i> .....	657
4. Does Germany Need a (New) Research Ethics for the Social Sciences?... <i>Claudia Oellers and Eva Wegner</i> .....	673
<b>IV. FIELDS</b> .....	687
1. Migration and Demography .....	687
1.1 Migration and Globalization... <i>Martin Kahanec and Klaus F. Zimmermann</i> .....	689
1.2 Migration and Statistics... <i>Sonja Haug</i> .....	703
1.3 Internal Migration... <i>Andreas Farwick</i> .....	723
1.4 Fertility and Mortality Data for Germany... <i>Michaela Kreyenfeld and Rembrandt Scholz</i> .....	739
2. Measuring Competencies .....	753
2.1 Measuring Cognitive Ability... <i>Roland H. Grabner and Elsbeth Stern</i> .....	753
2.2 Measuring Cognitive Competencies... <i>Ulrich Trautwein</i> .....	769
2.3 Measuring Vocational Competencies... <i>Martin Baethge and Lena Arends</i> .....	783
2.4 Measuring Social Competencies... <i>Ingrid Schoon</i> .....	797
2.5 Subjective Indicators... <i>Beatrice Rammstedt</i> .....	813
3. Education and Research .....	825
3.1 Education Across the Life Course... <i>Hans-Peter Blossfeld</i> .....	825

3.2	Preschool Education... <i>C. Katharina Spieß</i> .....	841
3.3	Data in the Domain of Secondary School Education.... <i>Petra Stanat and Hans Döbert</i> .....	859
3.4	Knowing More about Vocational Training... <i>Steffen Hillmert</i> .....	877
3.5	Higher Education... <i>Andrä Wolter</i> .....	891
3.6	Adult Education and Lifelong Learning... <i>Corinna Kleinert and Britta Matthes</i> .....	911
3.7	Research, Science, Development... <i>Stefan Hornbostel</i> .....	929
4.	Economy and Labor Markets.....	943
4.1	Data from the Federal Employment Agency... <i>Stefan Bender and Joachim Möller</i> .....	943
4.2	More and Better Data for Labor Market Research. Proposals for Efficient Access to the Currently Unused Potential of Official Statistical Data... <i>Hilmar Schneider</i> .....	959
4.3	Interdisciplinary Longitudinal Surveys. Linking Individual Data to Organizational Data in Life-Course Analysis... <i>Stefan Liebig</i> .....	971
4.4	Organizational Data... <i>Stefan Liebig</i> .....	985
4.5	Firm-Level Data... <i>Joachim Wagner</i> .....	1001
5.	State, Family, and Health.....	1015
5.1	Public Finance... <i>Thiess Büttner</i> .....	1015
5.2	Household Income, Poverty, and Wealth... <i>Richard Hauser</i> ...	1027
5.3	Family Research... <i>Johannes Huinink</i> .....	1041
5.4	Intergenerational Relationships... <i>Bernhard Nauck and Anja Steinbach</i> .....	1057
5.5	Administrative Data from Germany's Statutory Health Insurance Providers for Social, Economic, and Medical Research... <i>Günther Heller</i> .....	1081
5.6	Provision for Old Age: National and International Survey Data to Support Research and Policy on Aging... <i>Hendrik Jürges</i> .....	1093
5.7	Income Provisions and Retirement in Old Age... <i>Tatjana Mika, Uwe Rehfeld and Michael Stegmann</i> .....	1107

6. Political and Cultural Participation and the Role of the Media .....	1123
6.1 Political Participation – National Election Study... <i>Rüdiger Schmitt-Beck</i> .....	1123
6.2 Civil Society... <i>Mareike Alscher and Eckart Priller</i> .....	1139
6.3 Culture... <i>Jörg Rössel and Gunnar Otte</i> .....	1153
6.4 Mass Media Research... <i>Heiner Meulemann and Jörg Hagenah</i> .....	1173
6.5 Judicature... <i>Wolfgang Heinz</i> .....	1197
6.6 Environment... <i>Cornelia Ohl and Bernd Hansjürgens</i> .....	1217
<b>TERMINOLOGY</b> .....	1231



### ***III. DATA PROTECTION AND RESEARCH ETHICS***



# **1. Data Protection and Statistics – A Dynamic and Tension-Filled Relationship**

*Peter Schaar*

Contact:

Peter Schaar  
The Federal Commissioner for Data Protection and Freedom of Information  
Der Bundesbeauftragte für den Datenschutz und die Informationsfreiheit  
Husarenstraße 30  
53117 Bonn  
Germany  
e-mail: [peter.schaar\[at\]bfdi.bund.de](mailto:peter.schaar[at]bfdi.bund.de)



## Abstract

New statistical methods have been developed for long-term storage of microdata. These methods must comply, however, with the fundamental right to informational self-determination and the legal regulations imposed by the Federal Constitutional Court. Thus, it is crucial to develop effective and coherent methods for protecting personal data collected for statistical purposes.

Recent decisions by the Federal Constitutional Court are likely to result in the outlawing of comprehensive, permanent statistical compilations comprised of microdata from a wide range of sources that are updated regularly. However, aside from such comprehensive methods, there are certainly other ways of using microdata that cannot be dismissed from the outset as violating constitutional legal norms.

Internet access to statistical microdata is likely to take on increased importance for scientific research in the near future. Yet this would radically change the entire landscape of data protection: the vast amount of additional information now available on the Internet makes it almost impossible to judge whether individuals can be rendered identifiable. In view of this almost unlimited information, individual data can only be offered over the Internet if the absolute anonymity of the data can be guaranteed.

Keywords: right to informational self-determination, census ruling of December 15, 1983, longer-term storage of microdata, primary statistics, secondary statistics, statistical confidentiality, absolute anonymization, de facto anonymization, additional information, pseudonymization, personal data profiles

## 1. Introduction

Statistics traditionally deals with the collection and evaluation of data on the personal or material situations of a large number of individuals or organizations: "Statistics means ... activity aimed both at measuring mass phenomena, combining the data into groups and publishing them."<sup>1</sup> A more recent textbook contains the following definition, "Statistics is a combination of mathematical methods used to assess mass phenomena. The data collected serves to describe the environment numerically and/or in the event of uncertainty to use this data as a decision-making aid."<sup>2</sup>

The purpose of data protection is, according to the Federal Data Protection Act, "to protect the individual against impairment of his right to

---

1 Meyers Konversationslexikon (Meyer's Conversational Encyclopedia) (1907), volume 18, under the term "Statistik".

2 Bückner, R. (1997): Statistik für Wirtschaftswissenschaftler (Statistics for Economists), 3rd Edition, 11.

privacy through the handling of his personal data.”<sup>3</sup> “Personal data means any information concerning the personal or material circumstances of an identified or identifiable individual (the data subject).”<sup>4</sup>

If and when statistics are used merely to evaluate information relating to institutions (government agencies, companies) or natural phenomena (e.g., weather), data protection issues are irrelevant. However, the situation is much more complex with data on personal circumstances, as the heated debate on the 1983/87 census showed. This “individual data” is linked to the data subject at least during the data collection phase, and may also involve personal data. Only in the course of further data processing and evaluation is the personal reference eliminated partially or completely. In the final analysis, data may only be published if it can be ruled out in all likelihood that conclusions can be drawn about individuals. Personal data are therefore rendered (de facto) anonymous within the framework of traditional statistics.

Data protection requirements are changing as new statistical methods focused on long-term storage of individual data (microdata) become available (especially in the context of longitudinal studies). This means the described method of data collection – in which the data are rendered anonymous, preventing access to personal data and publishing only aggregated results – is no longer adequate under all circumstances, and thus no longer practicable.

## 2. The right to informational self-determination

With its census ruling of December 15, 1983, the Federal Constitutional Court (BVerfG, *Bundesverfassungsgericht*) formulated the “fundamental right to informational self-determination”:

“Under the terms of modern data processing, the protection of the individual against the unlimited collection, storage, use and passing on of his/her personal data comes under the general right to free development of one’s personality set forth in Art. 2 Abs. 1 of the Basic Law in conjunction with Art. 1 Abs. 1 of the Basic Law [inviolability of human dignity]. The basic right warrants [...] the capacity of the individual to determine in principle the disclosure and use of his/her personal data.”<sup>5</sup>

The ruling defined the requirements that need to be met in order to ensure that personal data are processed in accordance with the German Constitution (*Grundgesetz*). It ruled that any risks of misuse must be taken into account

---

3 § 1 Abs. 1 BDSG (*Bundesdatenschutzgesetz*). The citations to German legal sources have been left in German to guarantee accuracy.

4 § 3 Abs. 1 BDSG.

5 BVerfGE [decision] 65, 1, 1.

when data are processed. Even data that may seem irrelevant in an isolated context has the capacity to become relevant in a different context (by data fusion and data matching). The Federal Constitutional Court hence ruled that “considering the fact that individual data can be stored without any technical restraint with the help of automatic data processing ... there is no longer any such thing as irrelevant data.”<sup>6</sup> According to the German Constitution, the collection and processing of data must be justified by reasons of compelling public interest; the prerequisites and scope of data processing must be regulated comprehensibly for citizens, and the principle of proportionality must apply.

Last but by no means least, the further processing of data must, in principle, be limited to the purpose for which it was originally collected, which is particularly relevant for the collection of data for statistical purposes. Contrary to the collection of personal data for a specific administrative task, the need to collect data for statistical purposes can only be described in abstract terms, as the results can and indeed should be used for multiple purposes. It is hence all the more important to ensure that statistical data processing is separated strictly from the processing of data for administrative tasks. The envisaged use of data to correct information in the identity register in the 1983 census was the main reason for the negative ruling by the Federal Constitutional Court.<sup>7</sup>

One of the major risks in terms of data protection is that personal data profiles can emerge that are capable of presenting a complete picture of an individual. Any such personality profiles are incompatible with the Basic Law. The Federal Constitutional Court already established this in 1969, in its Microcensus ruling on personality profiles:

“It would be incompatible with human dignity if the state claimed the right to register and catalogue people in their whole personality coercively, even if the data collected in a statistical survey was rendered anonymous, as this would treat people like objects that are accessible for data collection in every respect.”<sup>8</sup>

Pursuant to this case ruling, it is now compulsory to protect personal data collected for statistical purposes in an effective and coherent manner. As such, it is important that the measures taken be oriented to the concrete threat situation and take the risks associated with rapid technological advancement into account.

---

6 BVerfGE 65, 1, 45.

7 BVerfGE 65, 1, 63.

8 BVerfGE 27, 1, 6.

### 3. Technological change

The main regulatory approaches to data protection originate from the time of mainframe computers, when electronic data processing took place at remote computing centers in accordance with rigid principles. Storage units the size of refrigerators, punch cards, and continuous printing paper dominated the scene when the Federal Constitutional Court issued its census decision in 1983.

Three main changes have taken place that are important in this context: the dramatic increase in storage capacities, the flexible evaluation possibilities, even of huge databases (“data mining”), and the “liberation” of computers from computing centers, offering 24/7 access to databases via networks, particularly the Internet.

In view of these trends, certain protection concepts that date back to the 1980s and 1990s are no longer realistic in today’s world. This applies, for instance, to the approach of physically sealing off the use of statistical data processing from processing for other purposes. Nowadays, statistical data can, of course, be processed on separate systems.

When data users are to be offered the benefits of computer technology, it is virtually impossible to do so without giving them electronic access to statistical data – e.g., through networks. It is difficult, if not impossible, to explain to users in the scientific and political communities why they are confined to rigid evaluations in the form of statistical aggregates and why they are denied access to microdata. After all, it is precisely microdata that offer a wide range of opportunities for obtaining new information. Nonetheless, the risks associated with these convenient types of use must be considered carefully. If data are processed outside the “walls” of statistical offices, it is virtually impossible to control how it is used – for instance, whether it is being used in combination with other databases.

What is needed are concepts that develop new, flexible possibilities for utilization that meet the expectations of data users and that simultaneously safeguard effective, modern data protection.

#### *Statistical confidentiality as a special data protection requirement*

When developing data protection measures, it is crucial that the various legal, organizational, and technical measures are well coordinated. As such, the starting point is the obligation to maintain the statistical confidentiality, which aims first and foremost at ensuring – like all other regulations

governing secrecy<sup>9</sup> – that only authorised “insiders” have access to the data and that the data are safe from use by unauthorised persons. The regulations governing the obligation to maintain statistical confidentiality represent special data protection regulations that override the general data protection legislation. They are intended not only to take the special sensitivity of the respective data into account, but also to build trust between the data subject and those who collect the data, who are obligated to maintain the confidentiality of statistical data on individuals without the individual having to fear negative consequences. According to the Law on Statistics for Federal Purposes (BStatG, *Bundesstatistikgesetz*):

“Individual data on personal circumstances or the material situation provided for federal statistics shall not be disclosed by the incumbents and the persons specially sworn in to public service who are entrusted with the operation of federal statistics, unless otherwise stipulated by a special legal provision.”<sup>10</sup>

In principle, personal data may only be used for certain tasks defined by law. It is prohibited and a punishable offence to use data for any purposes other than those expressly permitted by law. The same applies to passing on data to third parties outside the respective area. However, the principle of purpose limitation does not apply to statistical results that do not contain any personal reference. Individual statistical data may also be used for scientific purposes under certain circumstances:

“For the purpose of scientific projects, the Federal Statistical Office and the Statistical Offices of the German *Länder* may transfer microdata to institutions of higher education or other institutions entrusted with tasks of independent scientific research if an allocation of the individual data are possible only with an excessive amount of time, expenses and manpower, and if the recipients are elected officials, persons specially sworn in for public service, or persons obligated according to Abs. 7.”<sup>11</sup>

Contrary to this exemption for scientific purposes, the BStatG does not contain any explicit obligation to render individual data anonymous when these data are stored at statistical offices; however, this obligation arises implicitly from the jurisprudence of the Federal Constitutional Court, particularly in its census ruling. The legislator took these terms of reference into account by issuing detailed regulations on the rendering anonymous of data in a large number of individual statistical regulations on the deletion of calculation input features. After all, § 10 of the BStatG defines certain minimum (albeit merely geographical) requirements specifying precisely what individual data can be stored for extended periods by saying exactly what is prohibited and by prohibiting the use of precise address details. Finally, § 21 of the BStatG

---

9 Examples: the duty to treat medical records confidentially, confessional secrets, secrecy of postpostal and telecommunications secrecy.

10 § 16 Abs. 1 Satz 1 BStatG.

11 § 16 Abs. 6 BStatG.

stipulates that it is prohibited to match individual data from federal statistics or to combine any such individual data with other information:

“It is prohibited to match individual data from federal statistics or to combine such individual data with other information for establishing a reference to persons, enterprises, establishments or local units for other than the statistical purposes of this Law or of a legal provision ordering a federal statistics.”

#### **4. Statistical methods and data protection**

Even though traditional statistics are based, by and large, on data that refer to individual survey units, they do not rely on having retroactive access to individual data – with the exception of checks carried out during the data collection and data processing phase (to ensure the data are complete, plausible and, to a limited extent, correct). Rather, statistics involve aggregates, namely numerical values that are analysed and matched, with comparison and evaluation of any changes in this data over time. In principle, any such aggregates do not contain personal data unless it is possible to trace the results back to persons indirectly. It may, for instance, be possible to draw indirect inferences about individuals from statistical results if the respective table cell relates to a small number of people. The same applies to special characteristic values – for instance, if all members of a group have the same characteristic values.

Statistics are not matched at the case level. Only when statistics are published must it be ensured that none of the above-mentioned scenarios occur and that relevant countermeasures are undertaken (for instance, combining survey units, less distinctive characteristic values). As a rule, the loss of information associated with this rendering anonymous of data does not have any serious consequences and can certainly be tolerated (as long as different tabulations are not restricted, so to allow flexible tabulations).

The further development of statistical methods has led to heightened data protection requirements. The evaluation of statistical aggregates is supplemented by a more detailed analysis of patterns of individual statistical units at the so-called “micro-level.” Group patterns are traced back to patterns in the lives of individuals, who may have been observed over an extended period of time. To this end, data on the individual need to be collected and, if applicable, matched over time (into a longitudinal dataset). The annual Microcensus surveys that are carried out on the households under review at regular intervals over four consecutive years are based on this model.

Generally speaking, these new methods involve microdata that can be linked multifunctionally and can be evaluated over time (as, for example, in

clinical studies). There are numerous ways of accessing so-called “microdata files” – for example, through personal references in case scenarios in which the data are linked by a general personal identifier that can be used in a wide range of surveys.<sup>12</sup> However, there is no doubt that any such personal identifier is incompatible with the above-mentioned requirements of the Federal Constitutional Court in Germany.<sup>13</sup> For this reason, the court is likely to declare comprehensive, permanent microdata statistics comprising regularly updated data from a wide range of sources to be unlawful. However, aside from these comprehensive methods, there are certainly ways of using microdata that cannot be dismissed from the outset as violating the constitutional requirements.

## 5. Measures aimed at safeguarding data protection

It goes without saying that the traditional method of rendering data anonymous and deleting individual statistics based on a type of “stage model” is not compatible with a method that links microdata. It may be possible to trace the individual data back to the data subject even at the micro-level in the long term, which means the data do actually represent personal data in the majority of cases.

As such, one very interesting option would be to randomly link data collected within the framework of completely different statistical surveys in order to gain new information. In addition to the additional informative value such a method would yield, another argument in favor of it is the flexibility of the results it would generate.

In terms of data protection, any such method would involve major risks, given the apparent difficulty – if not impossibility – of rendering data anonymous in order to prevent inferences being drawn about identifiable statistical units. This risk could be mitigated by effectively ensuring that the data are protected against unauthorised access. However, whether this could achieve adequate protection is questionable, at least where particularly comprehensive or diverse microdata files containing personal features are concerned.

The type and size of the database is important when it comes to gauging the risk of abuse. Generally speaking, it can be said that the more comprehensive the database and the more sensitive the data, the greater the risk. This explains why censuses (which cover full populations) must be rated differ-

12 Lenz, R. and Zwick, M. (2005): Integrierte Mikrodatenfiles – Methoden zur Verknüpfung von Einzeldaten (integrated microdata files – methods of linking individual data). In: Statistisches Bundesamt (Ed.): Statistik und Wissenschaft 10, 100.

13 See. n. 6, BVerfGE 27, 1, 6.

ently than surveys in which small random samples of an entire population are taken. Data abuse also occurs with random sampling, albeit only to the extent of the sampling units included. Thus, the “abuser” needs to know who is included in the survey.<sup>14</sup>

It is also important to distinguish between primary and secondary statistics. It is not possible to state simply which of the two methods is more data protection-friendly. Occasionally, it is claimed that primary statistics, which collect data on the data subject are much more intrusive than secondary statistics, which do not collect any “new” data. This is only partly true. With secondary statistics, data are used and linked that were generally collected for another purpose altogether. This explains why most secondary files go hand in hand with data being used for a different purpose. Besides, the data subjects are “unaware” of the fact that their data are being used. Thus, they never gave informed consent. They are hence unable to check whether the data collection is lawful, and are unable to influence the process. In data protection terms, reference is made to deviation from the “Principle of Primary Collection” (ethical principle of informed consent). After all, more comprehensive secondary statistics – for instance, the census envisaged for 2011 – presupposes that it is possible to link data from different sources in which a particular type of infrastructure is needed. The question must therefore be raised how it can be prevented that infrastructures set up to collect statistics can also be used to link databases outside of statistics, with potentially far-reaching consequences for the data subject.

### *5.1 Rendering persons anonymous: absolute or de facto?*

During the census debate of the 1980s, the most important question raised was: when do data lose their personal reference and when are they deemed anonymous or at least “de facto anonymous”?<sup>15</sup> Only data that are completely anonymous contain no personal reference whatsoever, and therefore do not come under data protection regulations, whereas with de facto anonymous data it cannot be ruled out that the personal reference can be made/restored if relevant additional information is available. Additional information describes the information needed to identify a person even if neither the person’s name nor any other direct personal data (e.g., telephone number) can be linked

---

14 It is easy to make this impossible by deleting some cases that were sampled from the file that is available for analysis.

15 Fischer-Hübner, S. (1986): Zur Anonymität und Reidentifizierbarkeit statistischer Daten (Anonymity and reidentifiability of statistical data). In: Mitteilungen des Fachbereichs Informatik der Universität Hamburg 143; Brunnstein, K. (1987): Über die Möglichkeit der Re-Identifikation von Personen aus Volkszählungsdaten (The possibility of reidentifying persons from census data). In: Appel, R. (Ed.): Vorsicht Volkszählung! 2nd Edition, Cologne.



with other information that uniquely identifies the person. With individual statistics, it is possible to restore the personal reference if certain characteristic values are disclosed and if these characteristic values can be associated with the data subject. As such it must be borne in mind that the boundaries between personal and anonymous data are fading in view of the rapid increase in data volumes, as ever more powerful computers are making it easier and easier to restore the personal reference retroactively.<sup>16</sup>

With fully anonymous data, there is no case scenario in which third parties can associate data with a person. Complete or genuine anonymization hence means that personal data is altered in such a way as to ensure that the data can no longer be assigned to the person (even if there is additional information available). Only data that have been rendered fully anonymous contain no personal references whatsoever.

According to the definition of the Federal Data Protection Act, “rendering anonymous” means the modification of personal data so that the information concerning personal or material circumstances can no longer be attributed to an identified or identifiable individual, or only with a disproportionate amount of time, expense or labour.<sup>17</sup> This statutory definition is confined to rendering data de facto anonymous. Pursuant to § 3a of the BDSG, the data controller must implement data reduction and data economy measures. Pursuant to § 3a (2) of the BDSG, use is to be made of the possibilities for aliasing and rendering persons anonymous, insofar as this is possible and the effort involved is reasonable in relation to the desired level of protection to be achieved. This also applies to statistics.

## 5.2 Pseudonyms as an expedient?

The use of pseudonyms is appropriate in cases where it is necessary to identify a person but where an assumed identity is sufficient, namely when the real personal details do not need to be known and when, on the other hand, (actual or absolute) rendering anonymous is not possible. This type of case scenario frequently arises in statistics if data are stored at the micro-level (for instance for longitudinal analyses).

Aliasing means replacing a person’s name and other identifying characteristics (e.g., name, account number or personnel number) with a label in order to preclude identification of the data subject or to render such identification substantially difficult.<sup>18</sup> Data stored under an alias generally contains

---

16 Mattern, F. and Langheinrich, M. (2001): Allgegenwärtigkeit des Computers – Datenschutz in einer Welt intelligenter Alltagsdinge (Omnipresence of computers – data protection in a world of intelligent everyday objects), 13.

17 § 3 Abs. 6 BDSG.

18 § 3 Abs. 6a BDSG.

some kind of personal reference – albeit indirect. As such, it is important to distinguish between various types of aliases that are used in various contexts:

With reference aliasing, the allocation feature is assigned to the data subject using a reference list (or reference file). Reference aliases can always be deleted by the data controller using the reference list. With disposable aliases, the assignment features are derived from personal data using special computing functions (hash functions). The methods used must be selected to ensure that inferences cannot be drawn from the result about the individual persons or the identification features used. Disposable aliases are particularly suitable for longitudinal analyses in scientific research projects and statistics. With this type of aliasing, however, the data stored under the alias can only be assigned to the person using the alias if the original data used to create the alias is known and if the attacker knows how the alias was created (“Brute-Force Attack”).

### 5.3 Research Data Centers

The Research Data Centers run by the Federal Statistical Office, the Statistical Offices of the German *Länder*, the Institute for Employment Research (IAB, *Institut für Arbeitsmarkt- und Berufsforschung*) within the Federal Employment Agency (BA, *Bundesagentur für Arbeit*), and the Statistics of the German Pension Insurance (RV, *Deutsche Rentenversicherung*), have for a number of years made an attempt to balance data protection requirements against the interests of the scientific community in using microdata. The statistical offices give scientists direct access to individual data, observing general data protection requirements.

The Research Data Centers focus on microdata that have been cleared for remote data access.

However, the scientists do not access the statistical raw data or individual data managed by the offices directly, they access micro datasets, so-called Scientific Use Files (SUFs) generated for various purposes in which only virtually or fully anonymous data are stored.

SUFs can be accessed on-site or off-site. With on-site access, the data are accessed in the protected facilities of the Research Data Centers, whereas with off-site use, the data are accessed outside the Research Data Center for a specifically defined research project.

As the statistical offices have no way of ensuring the data are used properly, extreme caution must be taken when rendering data files anonymous, taking all of the additional information available to science into account. Access to official individual data are hence subject to the provisions set forth in the Law on Statistics for Federal Purposes.

Intensive use is being made of the newly established Research Data Centers.<sup>19</sup> Yet this is certainly not where developments will end, as there continues to be a keen interest in making the utilization of data even more flexible and above all in facilitating access from any location. Access via the Internet will likely be of key importance in the future. However, this would change the whole environment in terms of data protection, as it would no longer be possible to estimate the additional knowledge that might have been used to render individuals anonymous. In view of the unlimited amount of additional information available, individual data can hence only be used for uncontrolled Internet access if their absolute anonymity can be guaranteed. Anonymity “of a lesser quality” is not sufficient in view of the unlimited possibilities that exist for linking the widest range of databases.

---

19 Federal Data Protection Commissioner, TB 21, no. 7.6.



## 2. Record Linkage from the Perspective of Data Protection

*Rainer Metschke<sup>\*†</sup>*

---

\* This advisory report was translated for consistency, however, only the German version has been authorized by the author.

† Rainer Metschke passed away in August 2010. He was Advisor for the Berlin Commission for Data Protection and Informational Freedom.



## Abstract

This article will discuss record linkage from the perspective of data protection. To begin with, it will examine data fusion, a prominent form of which is “statistical matching.” This procedure occurs anonymously and in this respect, hardly appears to be relevant for protection of information issues. This changes, however, when results released as scoring values can be linked back to the individuals from which they stem.

For data integration, the term “record linkage” refers to a condition in which data stemming from two or more collections but related to a single entity must be combined so as to yield a unitary dataset. This dataset then allows for, for example, inferences about individuals. Notably, however, the combining of individual entries from official statistics with other official or even general statistics so as to attempt to trace these back to any particular person is a criminal offence. Nonetheless, there do exist exceptions where penalties do not apply; the most notable of these are the laws pertaining to the Microcensus. Even so, the specificity of such statutory statistics essentially ensures that these data cannot be effectively combined by other parties. However, the informational and transparency obligations that the state imposes upon statistical authorities is said to erode the legal protection afforded to these authorities when combining data. With that, this erosion hinders the creation of any basis for constructing quality data.

In addition to discussing the aforementioned, this article presents a thought experiment. This experiment highlights a situation in which data depositaries, defined as third-parties independent of both researchers and statistics authorities, can be used to link statistics back to individuals in a legally feasible manner. Further, this article will offer an example of how legal protection of relevant parties may be maintained whilst still allowing for the combining of data for research purposes.

**Keywords:** data fusion, statistical matching, data integration, record linkage, reanalysis, census judgment, 2011 Census, reidentification ban, data depositary, statistical confidentiality

## Preliminary remarks

It is perhaps one of the most cherished dreams of an empirical social researcher to be able to link datasets from diverse surveys relating to individuals, families, and households. If, for example, three attributes are collected in a survey, the researcher can evaluate them by analyzing one attribute three times, two attributes three times, or all three attributes once. If it is possible to link this set of data with yet another set from an additional survey – one that also has three attributes – rather than yielding seven analytical combinations, there would be... etc., etc.

This is also the ideal of criminologists, intelligence agencies, credit agencies, personnel managers, and epidemiologists, among many others. In light

of current circumstances, one should not forget those who deal in address and bank account information. At the same time, this potential is the nightmare of civil liberties advocates, employee and personnel representatives, labor unionists, and those professions involved in data protection.

## **A brief introduction to data fusion**

At first glance, data fusion hardly seems relevant to the subject of data protection. “Statistical matching” functions through the anonymity of individual datasets, derived from diverse sources and whose data either do not, or only incidentally, describe one and the same person. Common attributes or variables are used as linking elements using probabilities. This results in synthetic and thus artificial people. However, in terms of data protection, the application of this technique can be dangerous. Should these data be used to develop scoring values with which one evaluates real people or households, there are potentially serious consequences. If the probability is high enough that belonging to a group of individuals leads to a particular pattern of behavior, the individual can hardly defend himself against it. Lawmakers have recently begun to recognize this as well. In this context, one might think of the planned prohibition of the use of genetic data in the signing of insurance and work contracts. In addition, the upcoming amendment to the German Federal Data Protection Act will purportedly include the obligation, when using scoring methods, to disclose the methods used and the economic value of the data to those who will be affected.

Because of the constitutional status that guarantees their freedom, science and research should not be affected by such regulations. However, such research institutions are in fact the source of these methods of data fusion and the resulting evaluations. The fruit of scientific research becomes common property through publication and, as such, becomes available for use.

## **Data integration**

A precondition for record linkage is the existence of identifying data or keys to code individuals, but also codes to unique attribute values or attribute combinations, such as unchanging medical values or DNA sequences. The content data on identical persons, households, or businesses from various sources are integrated, and the body of information expands.



## **Current situation and recommendations for data linkage in the 2001 KVI report**

In 2001, the KVI report estimated that the current situation regarding the linkage of statistical data was inadequate and that the only acceptable solution from the standpoint of data protection – that of explicit consent – was impracticable and inefficient. This conclusion was based on concerns about the potential non-agreement of respondents, even though the linked data can no longer be matched to individuals. This would infringe on the public right to independent testing of statistical microdata through reanalysis.

The 2001 KVI report proposed that the efficiency of statistical production be increased through the linking of microdata, for which a legal authorization would be required instead of the consent of the respondent. Exact data linking would then be considered legally harmless if it was undertaken in an isolated area with subsequent anonymization. In particular, primary data collections and process statistics could be integrated and, as a result, the respondents would be released from the statistical duty of disclosure.

## **An indirect modification of the reidentification ban in statistics for federal purposes**

With the penal provision in § 22 of the Law on Statistics for Federal Purposes (BstatG, *Bundesstatistikgesetz*),<sup>1</sup> the ban on reidentification of personal information from federal statistics under § 21 is enforced by penalty. § 21 of the BStatG forbids the merging of personal data from federal statistics with other federal statistics or general information for the production of an individual reference. For purposes in accordance with the BstatG, or for a one-time statistical legal ruling, the ban does not apply. It first comes into effect when the identifying data is at least separated from the personal information, if not completely deleted. There must therefore have been a process of factual anonymization undertaken in the isolated inner domain of the statistical office. In 1987, this prohibition was evaluated by the Federal Constitutional Court, who declared it to be a supplementary, confidence-building provision in the service of data protection.

On the other hand, if the stipulated statistical legal provisions imply tasks that permit data integration or even dictate it, this cannot be fundamentally classified as constitutionally dangerous. At the time of the 2001 KVI report, these regulations did not exist.

---

1 The citations to German legal sources have been left in German to guarantee accuracy.

In 2005, § 13a of the BStatG was rewritten as follows:

“Data records from statistics pursuant to § 13 Abs. 1, data from the statistical register, data specified in the Law on the Use of Administrative Data, and data obtained by the Federal Statistical Office and the Statistical Offices of the German *Länder* from generally accessible sources may be matched, provided that matching is required in order to obtain statistical information without conducting additional statistical surveys.”

The reference in § 13a BStatG limits the permitted data integration to economic and environmental statistics relating to corporations, businesses, and workplaces. The Federal Statistical Office and the Statistical Offices of the German *Länder* are making increasing use of these new possibilities.

### **Sensitive data security issues regarding the integration of personal data in official statistics**

In the German Microcensus resolution of 1969, the Federal Constitutional Court made a deliberation weighing the fundamental law guaranteeing inviolable areas of private life against the collective need for information from the individual citizen. This inviolable private domain escaped public authority. It is incompatible with human dignity, which must be guaranteed to all without restriction, for the state to reduce the person to a mere object and to mandate the registration and categorization of the entire private individual.

In their evaluation, the judges in the constitutional court stressed that not every statistical survey with the duty of disclosure is injurious to human dignity and the right to self-determination in this protected domain. As citizens bound to the collective, each person must accept the necessity of statistical collection within certain parameters and, what is more, as a precondition for systematic state governance.

It should be noted that the theory underlying this concept of individual rights to privacy and the first laws on data protection were developed in different spheres of constitutional law. In the census decision of 1983, the Federal Constitutional Court drew from the right to privacy in Art. 2 Abs. 1, in connection with the duty to preserve the inviolability of human dignity pursuant to Art. 1 Abs. 1, to create the foundational law for the protection of the individual against unrestricted collection, storage, use, and disclosure of his or her personal data. Particularly, under the conditions of modern data processing, state incursions into the right to “informational self-determination” are subject to an overriding collective interest, and must conform to the relevant norms as well as have a constitutionally legal foundation. An intervention must be of reasonable scope and must include technical and organizational means of protection.

The principle of the primacy of consent is common to almost all legal provisions surrounding data protection, which may only be restricted in cases of overwhelming collective interest, the principle of adhering to the strict purposes of collection and the imperative of data economy.

In this context, the census decision (VZU, *Volkszählungsurteil*) stipulates the following:

“It would be incompatible with informational self-determination if a legal order should permit a social order in which the citizens would no longer know who, what, and under what circumstances something is known about him. [...] This would not only compromise the individual opportunities for personal development, but also the common good, because self-determination is a basic functional requirement of a liberal democratic society founded on trade and the capacity of its citizens to participate” (VZU, 43.).

So, if the official statistical agencies intentionally link data on citizens from different sources without their knowledge or agreement, in actual fact behind their backs, for the portrayal of relatively large temporal and factual spheres of people’s lives, can this be constitutionally acceptable?

Once again, on the census decision:

“The data collection program may portray individual spheres of life, for example the citizen’s place of residence, but not his private information. Anything other than this would only be permissible if it were possible to unrestrictedly link the data collected with those maintained by administrative agencies, which maintain in part very sensitive datasets, or even to access such a data-sharing facility through a uniform personal ID or other identifier would be possible; a comprehensive registration and cataloging of personal data through the merging of individual circumstantial and personal data for the construction of personal data profiles of the citizenry is, even with the anonymity of statistical records, unacceptable” (VZU 65, 53.).

The specific nature of official statistics speaks against the permissibility of the exact matching of data from different sources. Duty of disclosure, enforceable by state law, is the rule – agreement to provide information tends to be the exception. There is no concrete purpose for the collection of data. New analyses are constantly being generated by data research centers in particular. Inflexible table programs are no longer the only modes of evaluation, making them easily controlled. Statistics is always data collection for the future. Frugality in data collection results only from limited resources for official statistics and from the efforts to downsize bureaucracy at the expense of official statistics.

This does not make the collection of official statistics inherently unconstitutional, however. To the contrary, the tendency described above is balanced by an official statistical policy of confidentiality with regard to personal data, which includes the imperative of keeping a strict separation of statistics from administrative processes, the resulting freedom from self-incrimination owing to the fulfillment of the duty of disclosure, the internal and external partitioning of the statistical offices, the earliest possible

division and deletion of auxiliary identifiers, and the wide-ranging obligation to inform the respondents of their legal rights.

To sum up these observations, weighing both sides according to the principle of “practical concordance,” it appears that the integration of individual statistical sets of personal data in this legal framework can result in an imbalance to the disadvantage of the individual citizen.

### **2011 Census: The largest integration of data in federal statistics**

As many are already aware, the pending law regulating the census through multi-stage matching of personal data is derived from an administrative process. This is supplemented by a very large sample as well as a census of buildings and housing as primary surveys, and is expected to fulfill the preconditions for the generation of census datasets that largely correspond to the demands of a classic population count. Geodata will be matched with address data, population information from the data stores of municipal registers, the Federal Employment Agency (BA, *Bundesagentur für Arbeit*), and from the public administration for civil servants and judges. By means of “statistical matching,” the data from the population sample is then used as a supplement in order to fill gaps in the data from administrative processes, or to carry out in-depth regionalized evaluations. A consistent personal identifier is not available by reason of constitutional law as outlined above. Thus, the matching will proceed through clearly identifiable data, such as names, birthdays, etc. Aside from the logistical problems associated with this approach, those in data protection are interested in whether this operation is one whose depth of intervention is acceptable, and whether the legal and technological-organizational securities to which statistical offices are bound could not somehow be reduced.

The majority of those entrusted with data protection see to it that the duties of constitutional law described in the preceding section are being fulfilled and the requirements of the statistical offices are being adhered to. On individual points, however, there is still disagreement.

It has hardly been discussed that the methods developed for the 2011 Census have given rise to instruments that, should they come into use in administrative processes, would no longer guarantee citizens freedom from disadvantage. One need only to observe the generation of households. The obligation to collect data primarily from the relevant individual – whether they are collected for the receipt of social services, the obligation to provide information, or for the preparation of an official document – has been countered for several years by the development toward widespread matching

of administrative data. However, proposed statutory regulations on the storage and preservation of data, as well as restrictions on intended use, are increasingly coming under the scrutiny of constitutional law. Already at the time of the census test legislation, our public agency had called for an expansion of statistical confidentiality to extend to and include methods of data integration developed by the statistical offices and commissioned by them. Fully developed and functioning methods of data collection and processing systems awaken new desires – one need only think of the “LKW-Maut” (German toll system for heavy commercial vehicles) that politicians are now seeking to use for personal vehicles.

### **Research with personal data without consent on the basis of legal privileges for non-statistical research**

Should research be undertaken with a single one-time authorization or sweeping research privileges using data without explicit agreement of the respondents, the following mandatory checks and authorization steps must be followed:

- The urgent necessity of the data for the research project must be established.
- Can the research project only be successfully carried out using the personal data, or also with pseudonymized or even only with anonymized data?
- Can anonymization or pseudonymization be carried out step by step?
- What agency will carry out the anonymization or pseudonymization (researchers, a neutral trusted third party, address procurement, data steward)?
- In research using primarily personal data, authorization from upper federal or state agencies, and frequently also from an informational or qualified data protection agency, is mandatory.

After extensive consultation and through the oversight of our agency over the last two decades, a three-digit number of research projects were successfully carried out, accompanied by consultation and testing, all based on the integration of longitudinal or cross-sectional data.

## **The boundaries of the research privilege pursuant to § 16 Abs. 6 of the Law on Statistics for Federal Purposes**

The dilemma with regard to the use of personal statistical data for scholarly research lies in the fact that gradual anonymization or pseudonymization in the research process is ruled out. The data are permitted to leave the sealed-off area of the statistical office only when they have been factually anonymized. Reidentification, as stated above, is forbidden under penalty of law (see above), and thus longitudinal and cross-sectional data integration is forbidden as well. Only within the statistical office is integration of personal data not completely unacceptable, as both the pending Census Regulation Law and the Microcensus Law show.

### **A data model with a data steward function**

#### *A hypothetical exercise – even today*

Take the case of a scientific research project in which the statistical probability and margin of error of the two datasets collected on the basis of the obligation to provide information need to be statistically verified: on the one hand, income data from the Microcensus, and on the other, data from the respective tax office (*Finanzamt*) – which soon may become the Federal Financial Supervisory Authority (BAFin, *Bundesanstalt für Finanzdienstleistungsaufsicht*).

What follows is a brief observation on the use of currently established identification numbers for the tax authorities. Observing the principles of the population census judgment of the Federal Constitutional Court, the regulations contained in 139b of the tax code stipulate that this individual personal identifier may only be permitted for the fulfillment of the legal tasks of the tax authorities. This pertains also for the use by other public and non-public offices. The legislators, however, built a loophole into the tax code with the formulation, “or a statutory provision that explicitly allows or imposes the collection of data and use of the identification number.” It is extremely doubtful, however, that this regulation conforms to constitutional law. Should such a proviso succeed in being brought off, it would have to be quickly brought to trial before the Federal Constitution Court. In everyone’s interest, as much for social and economic research as for the acceptance of official statistics, I would like to urgently advise against the demand for the use of tax identification numbers for the purposes of data integration outside of the legal tasks of the tax authorities. According to the wording of the law – and this underscores doubts about this regulation – it is even possible to

allow uses within a legitimate, government-sanctioned objective that are foreign to the intended purpose. And now back to the theme of this advisory report.

Statistical offices do of course receive personal data from the tax authorities – but anonymously. Strictly speaking, the quality of this data already excludes the possibility of data integration. How can the data on individuals from the Microcensus and tax authorities be clearly linked to individuals in reanalysis?

The statistical office pulls a sample from the address data of the current sample census with file numbers. The desired attribute data with file numbers are stored separately in a parallel file. The address data with birth year and file numbers are transmitted to a data steward, who determines the tax office responsible for the particular address and assigns a pseudonym to the file number. The data steward conveys the address data of the individuals, the birth year, and the pseudonym to the appropriate tax office so that no connection to the household may be identified. After the data steward has transmitted this information, the data containing addresses and birth years are deleted. Based on this transmitted information, the tax office determines the tax data and correlates these with a minimal margin of error to the pseudonym. The address data and birth year are then deleted by the tax office. The tax office then encrypts the tax data and hands it back to the data steward with pseudonyms. The data steward then tells the statistics office which file numbers have pseudonyms, without passing on the pseudonyms themselves. The statistical office then passes the personal data from the Microcensus on to the data steward coded in yet a different way, but with the file number clearly provided. The data steward is then not in a position to decipher the attribute data. He replaces the file numbers with pseudonyms and then merges the variously coded personal data from the sample census and tax offices. The corresponding pseudonym is then deleted. What remains is only an anonymous and arbitrary dataset number. Any existing data referring to the tax office and regional offices within the German states are also deleted at this stage. The data are then transferred to the researchers in this form. With the private keys exclusively given to them, the data content can only be read and evaluated in this factually anonymized context. Even greater anonymization arises as a result of deficiencies in matching at the various tax offices.

#### Problems:

The tax offices find out that data have been prepared on specific individuals for data integration. The fact that these are data from official statistics can be kept confidential by the data steward (these might also be data from the BA or other sources, as in the case of data integration around questions of

marginal employment). The allocation of the public and private keys to the tax and statistical offices should be conducted by the researchers themselves or by a second data steward. The public key can only be given by the first data steward to the data provider.

The statistical office learns indirectly which of their data have no exact correlations at the tax offices (are these just retirees?).

### **Is a legally acceptable framework for integration of statistical data to support scientific research conceivable?**

Among researchers, data can be classified as factually anonymized if the integration of the data content does not present any fundamental potential for further de-anonymization. In the domain of official statistics, the basic imperative of separation is not harmed. For data from preliminary data collections, this method may no longer be appropriate if it is not possible to reconnect separate auxiliary identifiers with survey characteristics using a file number, as per § 9 Abs. 2 BStatG.

Legal consideration is required to determine whether this method prohibits reidentification of the attribute data “for the creation of an individual reference” (§ 21 BStatG). This seems to me to be the case. The question remains open whether this is covered under the tasks delineated in the Law on Statistics for Federal Purposes (§ 21 BStatG). If a legal research contract is signed for the use of federal statistics and this includes the rights to “auxiliary functions” according to § 16 Abs. 6 BStatG, then this would be possible.

At this point in my reflections, it appears that the principles of the division of official statistics and administrative processes, as well as the mandatory freedom from self-incrimination for respondents, are not infringed upon in the construct discussed above. To legally reinforce these principles, it is essential to provide legal legitimacy for the legal entity of the data steward. It is widely recognized that conceptions of a privileged right to confidential information for research purposes are of little use.

If, however, there were some broad legal parameters established for the three functions of data stewards (anonymization or pseudonymization; the relevant linking function; and functions of data preservation, preparation, and archiving), the situation would look significantly different. Even today, if a data steward takes on these functions immediately as a lawyer or notary in the service of and with the knowledge of the respondents, the data would be subject not only to a pledge of confidentiality, but also to the right to refuse to give evidence and the prohibition of seizure.



The individual respondents in the hypothetical construct above know nothing of the task the data steward is carrying out to protect their informational self-determination. A data steward is not their “trusted confidant.” I argue that these gaps can be bridged by lawmakers. A main point of criticism about the efforts and claims for research privileges on the side of the federal government is that, with this, the legally privileged and protected domain will become vastly expanded. To the critics of this new “secrecy,” with their restrictions on the security and criminal departments, it may be replied that this does not produce any new domain that cannot be controlled by these agencies.

For the official statistics, an opening clause might be integrated into the Law on Statistics for Federal Purposes (§16 Abs. 6) dealing with data stewards or such institutions operating as public authorities under the direction of a notary (distinct from but comparable to a hospital under the direction of a physician).

To sum up, I propose using the hypothetical scenario outlined above as a model for a discussion that can be considered from numerous perspectives.

## **Prospects and perspectives on the model**

Assuming the above described legal parameters, the model can be integrated, either directly or in modified form, into a range of possible solutions for the problem as it is laid out in the following statements. It should be theoretically possible to analyze the representativity of voluntary surveys among persons who had previously participated in the Microcensus on a mandatory basis. In term of data protection laws, however, this should presuppose a linkage of compliance of those being surveyed with an initially strong pseudonymization and a concluding process of anonymization secured by a data steward. The analysis then yields data about statistical errors in the totals and aggregate of the voluntary surveys. There are, however, erasure deadlines for the Microcensus for voluntarily modifying things later on. The model can be assessed in this context too, for which is it highly suitable, as far as it concerns survey data with data from official registers such as the local population registry, which are very precise but not linked to or used to identify individuals.

With the integration of personal data of different dimensions, levels, and content, the model can provide an instrument that can be used at specifically designated points of crossing and linkage in order to secure anonymity at every level. This would be particularly useful in connection with work and personnel data from various sources.

## **A personal afterword on the hypothetical construct**

For political counsel based on complex analysis and not only for alleged failures, which can supposedly only be uncovered by reanalysis in the official statistical agencies, independent research projects such as that called for in the 1998 memorandum (“Prerequisites for success in empirical economic research and empirically based economic and social policy advice”) appear both necessary and promising. In my opinion, such projects will become possible with the involvement of the relevant data protection agencies following strict interdisciplinary assessments of their necessity. Also desirable for such projects would be something akin to the Swiss model of an open and accessible public dialogue on the content of these projects. If details were brought into the public discourse, however, the freedom from self-incrimination could be indirectly affected. I propose that this be discussed as well.

If Professors R. Hauser, G. G. Wagner, and K. F. Zimmermann had written their memorandum in English, the language that now appears to be compulsory, the 2001 KVI report and the German Data Forum (RatSWD) would probably not have developed so successfully thus far. For this reason, only the German version of this essay is authorized.

## **Recommendations**

Official statistics should determine, on the basis of the specific information needs of social and economic research, for which data stores data integration is recommended. Additionally, official statistics should highlight general legal, technical and methodological problems.

On the basis of the examples provided in this report, it should be determined whether, and with what necessary requirements, including legal requirements, the present constitutional barriers might remain untouched through the use of independent and legally chartered data stewards. This structure would allow for the use of data through the diverse possibilities available for pseudonymization and encoding.

The relationship between the Research Data Centers and data stewards is yet to be defined. Both must occupy different positions of responsibility from a data protection perspective in order to preserve the integrity of the ban against reidentification.

Recommendations for the modernization of statistical law for the creation of clear and fixed parameters to enable data integration should be set as the goal for an initial conceptual phase and a pilot project.

### 3. New Methods in Human Subjects Research: Do We Need a New Ethics?

*Karsten Weber*

Contact:

Karsten Weber  
Technische Universität Berlin  
Informatik und Gesellschaft  
Sekretariat FR 5-10  
Franklinstraße 28/29  
10587 Berlin  
Germany  
e-mail: kweber[at]cs.tu-berlin.de

## Abstract

Online surveys and interviews, the observations of chat rooms or online games, data mining, knowledge discovery in databases (KDD), collecting biomarkers, employing biometrics, using Radio Frequency Identification Device (RFID) technology – even as implants in the human body – and other related processes, all seem to be more promising, cheaper, faster, and comprehensive than conventional methods of human subjects research. But at the same time these new means of gathering information may pose powerful threats to privacy, autonomy, and informed consent. Online research, particularly involving children and minors as well as individuals belonging to other vulnerable groups such as ethnic or religious minorities, is in urgent need of an adequate research ethics that can provide reasonable and morally justified constraints for human subjects research. The paper at hand seeks to provide some clarification of these new means of information gathering and the challenges they present to moral concepts like privacy, autonomy, informed consent, beneficence, and justice. Some existing codes of conduct and ethical guidelines are examined to determine whether they provide answers to those challenges and/or whether they can be helpful in the development of principles and regulations governing human subjects research. Finally, some conclusions and recommendations are presented that can help in the task of formulating an adequate research ethics for human subjects research.

Keywords: human subjects research, online research, biomarkers, biometrics, autonomy, privacy, informed consent, research ethics

## 1. Preface

Social researchers today regularly use questionnaires, interviews, or observation in the conventional paper-and-pen style to gather data (including methods like Computer Aided Telephone Interviews (CATI)). Yet it seems more promising, cheaper, faster, and comprehensive to deploy new means of collecting information: online surveys and interviews, observations in chat rooms or online games (see papers in Hine 2005; Kaye and Johnson 1999; Lyons et al. 2005), data mining in online social networks like Facebook or on individual websites, knowledge discovery in databases (KDD, see Tavani 1999a; 1999b; Vedder 1999), collecting biomarkers such as tissue samples or hairs, employing biometrics like face recognition to identify persons (e.g., Clarke and Furnell 2005; Crosbie 2005), using RFID<sup>1</sup> technology to monitor,

---

1 In their rather technical article, Goodrum et al. (2006) provide an overview of what RFID technology (Radio Frequency Identification) is, how it works, and how it can be used. Roughly speaking, RFIDs are very small computer chips which can store and process information as well as receive and transmit data wirelessly across distances. In the case of passive devices without a power source, this distance ranges from a few centimeters to as

for instance, consumer behavior in supermarkets (e.g., Lockton and Rosenberg 2005), or even implanting RFIDs into the human body. (see EGE 2005).

With the exception of online surveys and interviews, as well as the implantation of RFIDs, these methods can be employed without the awareness and knowledge of those who are being scrutinized. These new techniques are urgently in need of an adequate research ethics. The ethical dimension is critical, for it is almost impossible to administer research completely by legal precepts. Moreover, this option is even not desirable since it implies interference with academic freedom, reduces flexibility, and delays research projects. In addition, national law cannot typically be applied to international research programs. Thus, a research ethics that is widely acceptable across national and cultural borders could potentially serve as a kind of soft law. It is important to stress that, while ethics can be effective in controlling human behavior in general and academic research in particular, it cannot replace law, something made clear by the recent data crime violations in Germany. If people are willing to break the law, laws will not inhibit them and ethics will not either. Nevertheless, the following discussion will deal with ethics rather than law, under the assumption that all means and actions deployed in human subjects research have met all legal requirements.

## 2. Research ethics

### 2.1 *Different understandings of research ethics*

In the search for an adequate research ethics one does not have to start from scratch. Certain organizations, institutions, and professional associations have already done valuable work in this area that will be referenced in this expert report (e.g., AAAS 1999; AoIR 2002). And yet it is clear that what is understood by research ethics is sometimes quite different in different cases, if one compares, for instance, Germany and the United States.

Research ethics in Germany (*Forschungsethik*) is often used in reference to principles first proposed by Max Weber (1904) or Robert K. Merton (1942). Merton's CUDOS scheme in particular is often cited, according to which science must fulfill the demands of: *communism* or *communalism* (results must be shared with the scientific community), *universality* (everybody shall be able to participate in science regardless of nationality, religion, culture, etc.), *disinterestedness* (scientists shall present results as if they had

---

far as one or two meters. In the case of active devices with a power source, the broadcasting range can be increased to around ten meters. RFIDs can be used as identification labels for products, animals, and even human beings; they often are mentioned as serious threats to privacy (e.g., Lockton and Rosenberg 2005).

no personal interest in their rejection or acceptance), *originality* (researchers shall aim to develop novel claims), and *skepticism* (science and its claims shall always be subject to critical examination).

Of course, these demands are widely accepted in the United States too, but they are complemented by principles and rules that guide daily research routines and the application of research methods. These principles and rules will be identified below. It is important also to note that German codes of conduct and ethical guidelines regarding social science or marketing research (ADM 2001; DGS/BDS 1992) already include some similar rules. However, these documents seem more concerned with the relationship of principal and agent rather than with the relationship of researcher and research subject – for instance in the DIN ISO 20252.

## 2.2 Principles in research ethics

It does not make sense to try to find moral rules specifically for guiding either online surveys, data mining, collecting biomarkers, or the application of RFIDs. Instead, it is important to identify more general principles that can be applied to all new techniques of gathering information. Since these principles are, of course, principles, it should not make a big difference whether they are being applied to conventional social science methods or to newer ones.

At the same time, principles are abstract in that they do not tell us which action to take in a certain situation. For instance, Immanuel Kant's moral imperative demands generalizability of reasons for taking a certain action but is silent with regard to morally acceptable actions. Thus, it is necessary to supplement principles with advice on how to implement them in the research process.

In its *Scientific Freedom, Responsibility and Law Program* (AAAS 1999), the American Association for the Advancement of Science identifies three basic principles to guide research on human subjects: autonomy, beneficence, and justice. This document also introduces supporting principles such as privacy and informed consent. Thus, in the first section of this article, these principles are described in general and then applied to new techniques of gathering information. This is followed by a short discussion and concludes with the presentation of some conclusions and recommendations. Due to a lack of space, it is impossible to provide a comprehensive discussion on the problem of new techniques in human subjects research. Therefore, this text focuses on some of the most pressing issues.

### 3. Basic principles and their application

In approaching the following discussion, it is important to note that the way that concepts like autonomy, beneficence, and privacy are understood is culturally determined. This does not necessarily imply moral relativism; however, the understanding advanced here is not the only possible and existing one. Nonetheless, this paper takes the position that respect for the following ethical principles should form a kind of default option in human subjects research. It is always possible to reduce the requirements that have to be respected, but a research ethics based upon universal human rights and dignity should not allow research that does not respect these principles. They can be understood as absolutes that can only be abandoned if, and only if, research subjects deliberately consent. Such a position makes it possible to adapt these principles to other cultural contexts without diminishing the core values of our own ethics.

#### 3.1 *Autonomy, informed consent, and privacy*

Whether a person is to be granted autonomy or is already, by virtue of being a person, autonomous, is a question that has been discussed at least since the beginnings of Greek philosophy. The debate over informed consent has a more recent twentieth-century history, particularly as it pertains to the ethics of medicine and bioethics (see Sade 2001). The significance of privacy has been formulated at least since the hallmark paper of Warren and Brandeis, “The Right to Privacy,” published in 1890.

##### 3.1.1 General remarks

The concept of autonomy is a versatile one that can be filled with diverse meanings. In general, one can use “autonomy” as a descriptive term as well as ascriptive term. As *descriptive* term, it “[...] refers to people’s actual condition and signifies the extent to which they are meaningfully ‘self-governed’ in a universe shaped by causal forces” (Fallon 1994: 877). To be autonomous, a person must meet certain criteria like being able to make decisions on rational grounds. Simultaneously, the term presupposes a set of conditions, such as the absence of coercion. Used as an *ascriptive* term, “[...] autonomy represents [...] their right to make and act on their own decisions, even if those decisions are ill-considered or substantively unwise” (ibid.: 878). It is important to stress that this understanding of autonomy focuses on the individual. At least in some non-Western societies it is the case that either some adults, frequently women, are not granted autonomy, or the idea that individuals should or do have the opportunity to make independent decisions



is essentially denied (see Olinger et al. 2005). Since autonomy in its descriptive sense is a matter of degree, it has often been argued even in Western societies that certain circumstances allow for interference with a person's individual decision; such a perspective is often called "paternalism" (see Scoccia 1990) and will be discussed with reference to beneficence below.

In order to make autonomous decisions, some conditions must be met; being informed is one of these basic requirements. But informed consent is not always required of human subjects research. Gathering information that is publicly accessible – for instance, the content of television and radio programs or conducting observation in public spaces – does not require consent (see ASSS 1999: 7). That means that the distinction between private and public sphere is extremely important to human subjects research. If a researcher interferes with a person's private sphere or privacy, informed consent must be obtained (see Jacobson 1999: 135).

The shortest definition of privacy probably was coined by Samuel D. Warren and Louis D. Brandeis in 1890, who defined privacy as "the right to be let alone." Although their definition was and remains influential, far more detailed theorizations of privacy have emerged in recent years. In the context of human subjects research, and particularly with regard to new techniques of information gathering, the "control theory" and the "restricted access theory" of privacy (Tavani 1999b) should be mentioned. In control theory "[...] one has privacy if and only if one has control over information about oneself." (ibid.: 267). According to restricted access theory, "[...] an individual has privacy in a 'situation' if in that particular situation the individual is 'protected from intrusion, interference, and information access by others'" (ibid.). It must be again stressed here that the notion of privacy, like that of autonomy, is culturally biased.<sup>2</sup>

### 3.1.2 Application

Although autonomy has been discussed for a much longer time than informed consent and privacy, the latter two seem to be more important for human subjects research ethics. Privacy and informed consent are necessary prerequisites of autonomy insofar as it concerns the application of the new methods of information gathering mentioned above.

To respect privacy, it is essential to develop at least a working definition of private and public spheres. For instance, there is currently an intense debate about whether web pages, chat rooms, Usenet forums, and the like are public or private spaces. Often web pages are compared to radio and tele-

---

2 See, for example, the papers in the journal *Ethics and Information Technology* 7 (1) 2005. Yet, with reference to Newell (1998) one can arguably deny that such differences really exist.

vision broadcasting, which are publicly accessible and therefore allowed to be scrutinized without asking for any kind of consent from the broadcaster (AAAS 1999: 7).

However, without further indications one cannot presume that the creation and publication of Internet web pages automatically implies consent to their use for research purposes. One indication, for example, that authors of web pages do not consent to certain types of research use is when their pages contain so-called meta-tags which say that the respective web page must not be included in the index of a search engine like Google.<sup>3</sup> According to the control/restricted access theories of privacy, the use of such technical strategies is a way that authors of web pages try to take control over the flow of information.

Additionally, a main difference between radio and television broadcasting on the one hand and web pages on the other, is that web pages regularly contain information directly related to identifiable persons. Since human subjects research must meet the requirement of beneficence (see below), collecting information from such web pages potentially can cause harm to their authors. Clearly, these are cases that call out for informed consent. Finally, gathering data from web pages might interfere with copyright and intellectual property rights, which would also make informed consent mandatory (see Allen et al. 2008; Berry 2004; Carusi 2008; Grimes 2008; Hudson and Bruckman 2004; Jacobson 1999).

Some of the arguments mentioned above imply that the information to be collected is publicly accessible and that public accessibility already offers necessary as well as sufficient criteria for abandoning the requirement of informed consent. Conclusively, if information is not publicly disclosed, consent that this information may be used in human subjects research cannot be taken for granted. Thus, new methods of gathering data, like observations in chat rooms or online games, data mining in online social networks, or knowledge discovery in databases, all demand that informed consent is acquired explicitly. Because access to such data is regularly restricted by passwords and other technical and organizational means, it cannot be presumed that these sources of data are supposed to be publicly accessible (see Tavani 1999a; 1999b; Vedder 1999). Rather, they must be understood as belonging to the private sphere of a certain group of individuals or a sub-culture. Consequently, to support autonomous decision making with regard to participation in human subjects research, it is vital to ask for consent.

With regard to other methods, like collecting biomarkers such as tissue samples or hairs, employing biometrics, or using RFID technology to monitor persons' behavior, researchers must regularly assume that those who are being observed conceive of their behavior as something that belongs to their private sphere. Typically, biomarkers are not intentionally, but rather

---

3 Or in cases the website includes a file called "robots.txt," again with specific entries.

accidentally, put into circulation. Therefore, one cannot assume consent has been given to further investigation in human subjects research. The collection of biomarkers in the context of medical treatment proves this as a general rule, for it is mandatory to ask explicitly for consent and to inform the individual of potential risks and consequences. Therefore, if human subjects research conducted outside contexts like medical treatment is at stake, particularly if information about health status or the consuming habits of individuals are being gathered, informed consent, from an ethical point of view, seems mandatory (see Bayertz et al. 2001; see also below on “Beneficence”).

Lastly, using RFID implants for research purposes seems entirely inappropriate. After implantation, subjects have virtually lost their ability to autonomously stop the research process. Simultaneously, the risks of scarring, infection, and other health risks are quite difficult to evaluate, particularly for non-specialists (EGE 2005: 18; see below). Thus, for the application of RFIDs, well informed consent is difficult or even impossible to obtain.

### 3.2 *Beneficence*

#### 3.2.1 General remarks

Generally speaking, beneficence as a moral claim means that with our actions we aim to promote the good of others and increase their benefits, and also try to prevent harm from others. As a guiding principle for our behavior, beneficence requires us to take the consequences of our actions into account. Therefore, it is necessary to try to forecast the possible and likely outcomes of current and future decisions. Obviously, such forecasts are often difficult or even impossible.<sup>4</sup> However, that is not the main problem posed by beneficence; rather, it is that beneficence may collide with the principle of autonomy. In moral as well as in political philosophy there is rigorous debate over whether the benefit to a person can be objectively measured or whether it can only be evaluated from an individual point of view. The latter view purports that, for instance, what is harmful for one person could be a benefit for another.

#### 3.2.2 Application

Human subjects research in general can expose individuals to certain risks of harm. Although it might be difficult or even impossible to define one single

---

4 One important response to this problem is represented by the “precautionary principle” (see Morris 2000). Except for the EGE Opinion No. 20 (EGE 2005: 17), the precautionary principle has not been explicitly taken into account in those codes of conduct or ethical guidelines referred to in the text at hand.

standard of good and harm that is acceptable for every person, it is obvious that some consequences of human subjects research are unambiguously intolerable: mental or physical harm, discrimination, damage or loss of property, and the like. Particularly where methods are employed that make subjects (potentially) identifiable, and thus may expose them to such consequences, it is extremely important to take the principle of beneficence into account. With regard to, for instance, data mining, KDD, biomarkers, or biometrics, the risk assessment of possible identification is therefore mandatory. It could be that a single set of data does not allow for identification of subjects, but a combination of several different databases would present this possibility. In such cases, research subjects must be informed and asked for their consent – regardless of whether publicly accessible data is used or not. Particularly if risk assessment is impossible or does not provide viable evidence, the principle of beneficence may even require a cessation of research.

### *3.3 Justice*

#### *3.3.1 General remarks*

The third basic principle that shall guide research is justice, which demands a fair distribution of risks and benefits resulting from our actions. As described in the report of the AAAS (1999: 3): “Since the fruits of knowledge can come at a cost to those participating in research [...] justice [...] seeks a fair distribution of the burdens and benefits associated with research, so that certain individuals or groups do not bear disproportionate risks while others reap the benefits.” In fact, justice can be interpreted as impartial beneficence. It is important to stress that a fair distribution of burdens and benefits does not necessarily imply equality but equity in distribution.

#### *3.3.2 Application*

With regard to human subjects research “[...] justice is perhaps the most elusive [principle] in terms of application and understanding” (AAAS 1999: 14). As already mentioned, it is quite difficult to make exact determinations around the notions of good, harm, and beneficence. If the term “justice” is understood as impartial beneficence, it is still unclear how benefits and burdens of human subjects research might be shared. In fact, it might be argued that since it is difficult to determine positive as well as negative outcomes of research, it would not make sense to talk about the just distribution of these outcomes. Nonetheless, there is one notion of justice that has a direct impact on human subjects research: the principle of justice does

not allow for the instrumentalization of individuals or groups of individuals who certainly will never be, not even potentially, beneficiaries of a specific research program.<sup>5</sup>

#### 4. A special problem

A very important question concerning autonomy, privacy, and informed consent is the problem of research on children and minors, for example, in the behavioral sciences, social sciences, epidemiology, or pedagogy and educational sciences. This kind of research is continuously growing as the critical importance of the first years of the life course becomes more and more obvious, for example, in research pertaining to school and preschool education. Yet, if one takes a closer look at the existing codes of ethics and codes of conduct as well as at the literature concerning ethics in general, children and minors are occasionally mentioned, but it is very difficult to find concrete advice for research.

For instance, in the report, *Ethical and Legal Aspects of Human Subjects Research on the Internet*, only two sentences on minors can be found: "For example, minors could respond to a study involving inappropriate materials for their age without the researcher's knowledge" (AAAS 1999: 8) and "Researchers are obligated by federal policies and professional ethics to provide special consideration for vulnerable members of the community, such as children and persons of diminished mental capacity" (ibid.: 5).

In the *ICC/ESOMAR International Code on Market and Social Research* one comes across statements like, "Market researchers shall take special care when carrying out research among children and young people" (ICC/ESOMAR 2007: 2), or "Researchers shall take special care when interviewing children and young people. The consent of the parent or responsible adult shall first be obtained before interviewing children" (ibid.: 6).

The *EGE Opinion No. 20* says that "ICT devices should be implanted in minors and legally incapacitated only if this is done in accordance with the principles set out in the Council of Europe Convention on Biomedicine and Human Rights" (EGE 2005: 31).

Finally, both the code of ethics adopted by the German Sociological Association (DGS/BDS 1992) as well as the 2001 report, "Standards for Quality Assurance for Online Surveys" (*Standards zur Qualitätssicherung für Online-Befragungen*) by the Working Group of German Market and

---

5 To use an example from biomedical research: pharmaceutical tests in developing countries which test drugs that will not be sold in the respective countries or that are not affordable for the research subjects themselves.

Social Research Institutes (ADM, *Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute*), are almost completely silent on the question of ethics in human subjects research. Furthermore, neither the ADM's "Guidelines on the Use of Mystery Research in Market and Social Research" (*Richtlinie für den Einsatz von Mystery Research*<sup>6</sup> in der Markt- und Sozialforschung) from 1995, nor its "Guidelines for Online Surveys" (*Richtlinie für Online-Befragungen*) from 2000 refer to children or minors at all.

To summarize, scholars will on the whole find few if any references to legal regulation, but even fewer instructions concerning the design of research on children and minors. However, one can find detailed recommendations in "Ethical Decision-Making and Internet Research", a document put out by the Association of Internet Researchers in 2002.<sup>7</sup> Unfortunately, one must derive these recommendations from three sample consent forms for parents and children involved in Internet research. Nevertheless, this might be a useful point of departure for considering this question in individual contexts.

## 5. Conclusions, requirements, and recommendations

The above-mentioned principles of autonomy, privacy, informed consent, beneficence, and justice are just as important as ethical principles themselves. As guiding principles for human subjects research, they have particular potency. However, they must be supplemented by rules for their application in research design and on research processes. So far, some deficits can be identified that point directly to some specific requirements and recommendations.

- Human subjects research programs should employ risk assessment procedures concerning the potential for identification of research subjects if multiple databases are combined;

---

6 "Mystery Research" includes covert observations in chat rooms and other similar online sites. Significantly, the ADM assumes that mystery research does not require informed consent. Nevertheless, it should be noted that the ADM does provide "Guidelines for Interviewing Minors" (*Richtlinie für die Befragung von Minderjährigen*) (ADM 1996) that contain comparable recommendations to those of the AoIR (2002). However, a detailed interpretation of all the ADM's guidelines would probably reveal some incompatibilities and even contradictions, for example with regard to informed consent in the case of online ("mystery") research on children and minors.

7 Association of Internet Researchers, for more information refer to <<http://www.aoir.org>>, last visited 01/05/2009.

- Thresholds must be defined concerning acceptable risks for research subjects that also differentiate between children, minors, and adults;<sup>8</sup>
- A more appropriate definition of beneficence must be developed that focuses on preventing individual harm. The goal of working for the good for each research subject is highly implausible, difficult, and perhaps even impossible to obtain;
- Specific and concrete rules concerning human subjects research on children and minors must be developed and then incorporated into codes of ethics and codes of conduct. Specific attention must be given to the issue of data collection that involves children and minors who are now adults, particularly with regard to panel surveys. In such situations it is recommended that research subjects be asked for the renewal of informed consent. In the case of a denial it would then be mandatory to delete all personal data, for example, names and addresses, out of respect for autonomy, privacy, and beneficence (sometimes it might even be necessary to consider to delete *all* existing data to comply with copyright and intellectual property rights);
- As far as possible, thresholds, definitions, and rules concerning human subjects research must not be based on particular, culturally determined customs and traditions. Reference to customs and traditions makes it more difficult to adopt a general research ethics to different cultural contexts;
- Because such definitions and thresholds are often difficult to generalize, particularly in case of long-term research projects – projects involving a very large number of participants, or projects involving subjects with greater vulnerability like children or members of ethnic or religious minorities (because, for instance, if such groups are small it might be easy to deanonymize data and identify individuals) – it might be necessary to establish ethics committees specifically for human subjects survey research, similar to those that exist in (bio-)medical research programs.

---

8 Such thresholds already exist in animal research but are probably not sufficient for human subjects research. Animal-related research does not deal with questions concerning autonomy, privacy, and informed consent. However, it might be helpful to study the history of how such thresholds developed.

## References:

- AAAS (1999): Ethical and Legal Aspects of Human Subjects Research on the Internet. American Association for the Advancement of Science. <http://www.aaas.org/spp/sfrl/projects/intres/report.pdf>. [Last visited 08/21/2008].
- ADM (1995): Richtlinie für die Befragung von Minderjährigen. Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e.V. [http://www.adm-ev.de/pdf/R05\\_D.pdf](http://www.adm-ev.de/pdf/R05_D.pdf). [Last visited 01/05/2009].
- ADM (1996): Richtlinie für den Einsatz von Mystery Research in der Markt- und Sozialforschung. Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e.V. [http://www.adm-ev.de/pdf/R02\\_D.pdf](http://www.adm-ev.de/pdf/R02_D.pdf). [Last visited 01/05/2009].
- ADM (2000): Richtlinie für Online-Befragungen. Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e.V. [http://www.adm-ev.de/pdf/R08\\_D\\_07\\_08.pdf](http://www.adm-ev.de/pdf/R08_D_07_08.pdf). [Last visited 01/05/2008].
- ADM (2001): Standards zur Qualitätssicherung für Online-Befragungen. Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e.V. [http://www.adm-ev.de/pdf/Onlinestandards\\_D.PDF](http://www.adm-ev.de/pdf/Onlinestandards_D.PDF). [Last visited 08/21/2008].
- Allen, G.N./Burk, D.L. and Ess, Ch. (2008): Ethical Approaches to Robotic Data Gathering in Academic Research. *International Journal of Internet Research Ethics* 1 (1), 9-36. [http://www.uwm.edu/Dept/SOIS/cipr/ijire/ijire\\_1.1\\_allen.pdf](http://www.uwm.edu/Dept/SOIS/cipr/ijire/ijire_1.1_allen.pdf). [Last visited 08/27/2008].
- AoIR (2002): Charles Ess & AoIR ethics working committee. Ethical decision-making and Internet research: Recommendations from the AoIR ethics working committee. <http://www.aoir.org/reports/ethics.pdf>. [Last visited 08/27/2008].
- Bayertz, K./Ach, J.S. and Paslack, R. (2001): Wissen mit Folgen. Zukunftsperspektiven und Regelungsbedarf der genetischen Diagnostik innerhalb und außerhalb der Humangenetik. *Jahrbuch für Wissenschaft und Ethik*, Band 6, 271-307.
- Berry, D.M. (2004): Internet research: privacy, ethics and alienation: an open source approach. *Internet Research* 14 (4), 323-332.
- Carusi, A. (2008): Data as Representation: Beyond Anonymity in E-Research Ethics. *International Journal of Internet Research Ethics* 1 (1), 37-65. [http://www.uwm.edu/Dept/SOIS/cipr/ijire/ijire\\_1.1\\_carusi.pdf](http://www.uwm.edu/Dept/SOIS/cipr/ijire/ijire_1.1_carusi.pdf). [Last visited 08/27/2008].
- Clarke, N. and Furnell, S. (2005): Biometrics – The promise versus the practice. *Computer Fraud & Security* 9, 12-16.
- Crosbie, M. (2005): Biometrics for enterprise security. *Network Security* 11, 4-8.
- DGS/BDS (1992): Ethik-Kodex der Deutschen Gesellschaft für Soziologie (DGS) und des Berufsverbandes Deutscher Soziologen (BDS). Berufsverband Deutscher Soziologinnen und Soziologen e.V. <http://www.bds-soz.de/images/stories/formulare/ethik.pdf>. [Last visited 08/21/2008].
- EGE (2005): European Group on Ethics in Science and New Technologies to the European Commission: Opinion on the ethical aspects of ICT implants in the human body. Opinion No. 20. [http://ec.europa.eu/european\\_group\\_ethics/docs/avis20compl\\_en.pdf](http://ec.europa.eu/european_group_ethics/docs/avis20compl_en.pdf). [Last visited 08/19/2008].
- Fallon, R.H. Jr. (1994): Two Senses of Autonomy. *Stanford Law Review* 46 (4), 875-905.
- Goodrum, P.M./McLaren, M.A. and Durfee, Adam (2006): The application of active radio frequency identification technology for tool tracking on construction job sites. *Automation in Construction* 15 (3), 292-302.



- Grimes, S.M. (2008): Researching the Researchers: Market Researchers, Child Subjects and the Problem of "Informed" Consent. *International Journal of Internet Research Ethics* 1 (1), 66-91. [http://www.uwm.edu/Dept/SOIS/cipr/ijire/ijire1.1\\_grimes.pdf](http://www.uwm.edu/Dept/SOIS/cipr/ijire/ijire1.1_grimes.pdf). [Last visited 08/27/2008].
- Hine, Ch. (Ed.) (2005.): *Virtual Methods. Issues in Social Research on the Internet*. Oxford/New York.
- Hudson, J.M. and Bruckman, A. (2004): "Go Away": Participant Objections to Being Studied and the Ethics of Chatroom Research. *The Information Society* 20, 127-139.
- ICC/ESOMAR (2007): ICC/ESOMAR International Code on Market and Social Research. International Chamber of Commerce. [http://www.iccwbo.org/uploadedFiles/ICC/policy/marketing/Statements/ICCESOMAR\\_Code\\_English.pdf](http://www.iccwbo.org/uploadedFiles/ICC/policy/marketing/Statements/ICCESOMAR_Code_English.pdf). [Last visited 08/29/2008].
- Jacobson, D. (1999): Doing Research in Cyberspace. *Field Methods* 11 (2), 127-145.
- Kaye, B.K. and Johnson, Th.J. (1999): Research Methodology: Taming the Cyber Frontier. *Techniques for Improving Online Surveys*. *Social Science Computer Review* 17 (3), 323-337.
- Lockton, V. and Rosenberg, R.S. (2005): RFID: The next serious threat to privacy. *Ethics and Information Technology* 7, 221-231.
- Lyons, A.C./Cude, B./Lawrence; F.C. and Gutter, M. (2005): Conducting Research Online: Challenges Facing Researchers in Family and Consumer Sciences. *Family and Consumer Sciences Research Journal* 33 (4), 341-356.
- Merton, R.K. (1942): Science and Technology in a Democratic Order. *Journal of Legal and Political Sociology* 1, 115-126. [Reprinted as "The Normative Structure of Science" in Merton, R.K. (1974): *The Sociology of Science*. Chicago].
- Morris, J. (Ed.) (2000): *Rethinking Risk and the Precautionary Principle*. Oxford.
- Newell, P.B. (1998): A Cross-Cultural Comparison of Privacy Definitions and Functions: A Systems Approach. *Journal of Environmental Psychology* 18, 357-371.
- Olinger, H.N./Britz, J.J. and Olivier, M.S. (2005): Western privacy and ubuntu: influence in the forthcoming data privacy bill. In: Brey, Ph./Grodzinsky, F./Introna, L. (Eds.): *Ethics of New Information Technology*. Enschede.
- Sade, R.M. (2001): Autonomy and Beneficence in an Information Age. *Health Care Analysis* 9, 247-254.
- Scoccia, D. (1990): Paternalism and the Respect for Autonomy. *Ethics* 100 (2), 318-334.
- Tavani, H.T. (1999a): Informational privacy, data mining, and the Internet. *Ethics and Information Technology* 1, 137-145.
- Tavani, H.T. (1999b): KDD, data mining, and the challenge for normative privacy. *Ethics and Information Technology* 1, 265-273.
- Vedder, A. (1999): KDD: The challenge to individualism. *Ethics and Information Technology* 1, 275-281.
- Warren, S.D. and Brandeis, L.D. (1890): The Right to Privacy. *Harvard Law Review* IV (5), 193ff.
- Weber, M. (1904): Die "Objektivität" sozialwissenschaftlicher und sozialpolitischer Erkenntnis. *Archiv für Sozialwissenschaft und Sozialpolitik* 19, 22-87.



#### 4. Does Germany Need a (New) Research Ethics for the Social Sciences?

*Claudia Oellers and Eva Wegner*

Contact:

Claudia Oellers  
Business Office of the German Data Forum (RatSWD)  
c/o DIW Berlin  
10108 Berlin  
Germany  
e-mail: coellers[at]ratswd.de

Eva Wegner  
University of Cape Town  
Department of Political Studies  
Rondebosch, 7700  
South Africa  
e-mail: wegner.eva[at]gmail.com

## **Abstract**

This paper evaluates the German, UK, and US approaches to dealing with research ethics in the social sciences. It focuses (1) on the extent to which these research ethics frameworks protect the key rights of research subjects and (2) the extent to which they take into account the methodology and approaches used in the social sciences and do not simply emulate those of the natural sciences.

The US approach represents a highly regulated and partly bureaucratic approach where the ethics review is modeled on the methodologies of the natural sciences. In the UK, in contrast, a social science research ethics framework has been developed that remedies some of these shortcomings. It is implemented through pressure from funding institutions and is designed to respond to the needs of social science research. The German social science ethics framework consists of non-binding codes of conduct, guidelines about good scientific practice, and ethics codes of the German professional associations and funding institutions. We find that ethical behavior in Germany is typically understood as ethical behavior towards peers. We recommend the establishment of a new research ethics framework for the social sciences in Germany that is modeled on the approach developed in the UK.

Keywords: research ethics, good scientific practice, institutional review boards

## **1. Introduction**

Several recent papers have addressed the need for rethinking research ethics in the social sciences (e.g., Lane 2009, Weber 2010). Two reasons are typically given special emphasis. The first is based on new forms of collaboration among social scientists and researchers from other fields that are judged ethically more sensitive, especially biomedical research. Research that looks, for instance, at the behavioral consequences of genetic configurations can easily confront social scientists with new ethical dilemmas. The second reason given is that technological developments now allow for a large amount of data to be exchanged through or are freely accessible on the Internet. Data that are either available from agencies or that citizens themselves make accessible – for instance, on their websites or in online forums – create new possibilities for data matching. This produces new challenges for an obviously fundamental principle in research ethics – the anonymity of research subjects. Indeed, these two developments are among the key motivations that gave rise to a new research ethics framework for the Economic and Social Research Council in the UK (see ESRC 2005).

These developments are certainly important and invite the reconsideration and revision of research ethics in the social sciences. At the same time, an exclusive focus on such new developments may bury the fact that the

existing framework for social science research ethics may already be inadequate for “standard” empirical work in the social sciences.<sup>1</sup> Our paper thus evaluates different ways of dealing with research ethics, focusing on two questions that must be at the core of any discussion and revision of social science research ethics. First, to what extent does a research ethics framework protect the key rights of research subjects, such as information and anonymity? Second, to what extent is that framework appropriate for social science research? That is, is it simply modeled on the natural sciences, or does it respond to the different methodologies and approaches used in the social sciences?

This report looks first at the German social science ethics framework, which is essentially one of non-binding codes of conduct, guidelines about good scientific practice, and ethics codes of the German professional associations and funding institutions. We find that ethical behavior is typically understood as ethical behavior towards one’s peers. Second, we discuss US and British approaches to research ethics in the social sciences.

The US approach can be seen as a highly regulated and relatively bureaucratic approach where the ethics review is modeled on the methodologies of natural sciences. The above-mentioned new framework used in the UK, in contrast, represents a reformist approach that is implemented through pressure from funding institutions and aims to respond to the needs of social science research. Finally, we address the question of whether either of these could serve as a role model for the social sciences in Germany.

## **2. The research ethics infrastructure in Germany**

In Germany, ethical requirements for research vary widely across research fields. Requirements are rigorous and legally binding in medical or biomedical research and much less so in the social sciences, where the only compulsory legal standard is the Data Protection Act (see Schaar 2010).

As in other countries, ethical questions in Germany have always been more prominent in the natural sciences than the social sciences. In the early 1970s, some universities had already established research ethics committees (RECs). In 1979 the German Medical Association, following an initiative of the German Research Foundation (DFG, *Deutsche Forschungsgemeinschaft*), recommended the introduction of RECs. In 1994, approval by a research ethics committee became compulsory for clinical trials following the Medi-

---

1 Whether these new trends do indeed constitute new challenges for social science is also contested. Greely (2008), for instance, argues that although many feel differently about it, the data commonly used in the social sciences is not less sensitive than information about health issues.

cial Product Act (MPG) and the fifth amendment to the Drugs Act (AMG). In 2004 a further amendment was enacted that implemented the good clinical practice directive of the EU (2001/20/EC). As a result, a majority of German medical faculties and medical research institutes now possess RECs.<sup>2</sup>

For the social sciences there is no comparable legal regulation for approval of research through a research ethics committee. No important funding institution or state agency has made it its mission or a priority to further, or to systematically address such standards in the social sciences. The only legal requirement to take into account in the social sciences is the federal law on data protection. This law addresses issues of consent, data gathering, storage, and processing for all kinds of research. It is an articulation of some general standards for data-related issues in scientific research, such as the duty to anonymize information.<sup>3</sup>

It is rather in the framework of professional self-regulation by professional associations of sociologists or psychologists that ethics questions are addressed in the social sciences in Germany. These professional associations have created ethics committees and established codes of ethics. For example, there is the joint code of ethics of the two professional associations of sociologists, the German Sociological Association (DGS, *Deutsche Gesellschaft für Soziologie*) and the Professional Association of German Sociologists (BDS, *Berufsverband Deutscher Soziologen*), which dates back to 1992. It addresses ethical standards in research – issues of integrity and objectivity as well as the protection of the research subjects – and also deals with relationships among academics, such as the duty of referees to state conflicts of interest. These two professional associations have established a joint ethics committee to which complaints on misconduct in all the areas covered by the ethics code can be brought. This ethics committee is supposed to help find consensual solutions, but it also has the prerogative to suggest sanctions, such as the temporary exclusion of a member or her full expulsion.

While this ethics committee may advise the professional associations on ethical questions, it is in no way involved with approving research projects from an ethical point of view. This type of ethical evaluation is performed, however, by the professional association of German psychologists, the German Psychological Society (DGPs, *Deutsche Gesellschaft für Psychologie*). Its ethics committee evaluates applications for which a funding institution has required a review. In this case, and for a fee, the ethics committee evaluates whether the goals and procedures of the project comply with

---

2 The tendency to consider ethical issues more important in biomedical research than in social sciences is also evident in the work of the National Ethics Council, established by the Federal Government in 2001 (since 2007, the German Ethics Council). The council has published reports and recommendations on several topics, but most of them concern the field of biomedical research.

3 For more details on legal requirements regarding data protection, see Schaar 2009.

ethical standards. A few social science departments in Germany, for example at the universities of Mannheim and Jena, have also established local ethics committees on their own initiative that review the research projects of faculty members.

More general standards for good scientific practice were defined in a set of guidelines established by the DFG, following a series of research misconduct cases in 1997 (DFG 1997). These guidelines encompass all fields of scientific research and focus strongly on questions of ethical behavior among researchers. The DFG recommended that universities establish their own principles on the basis of the DFG guidelines, and in 1998 decided that research institutions receiving funding from the DFG had to establish rules ensuring good scientific practice. In 1999, the DFG also created an institution – an ombudsmen committee – to investigate cases of scientific misconduct and to monitor the implementation of ethics guidelines.

After more than ten years it seems fair to say that the DFG guidelines have remained relatively inconsequential for promoting good scientific practice in research and teaching. Indeed, the reports of the ombudsmen themselves lament that there is little awareness of good practice and scientific misconduct.

Although German universities quickly adopted either the DFG rules or developed their own, they have made little effort to promote them.<sup>4</sup> By and large, researchers are unaware of the existence of these rules. This lack of awareness among researchers and the sorry efforts of universities to promote the rules was already pointed out by the DFG in 2001 and, according to the latest report in 2008, little progress has been made since. This report suggests that the awareness of good scientific practice could be increased via the implementation of another principle in the original guidelines (rule no. 2), namely making issues pertaining to good scientific conduct a standard item in the teaching and training of junior researchers. However, given that this suggestion has been largely ignored for the last ten years, it seems questionable whether this suggestion will have much effect.

The implementation of ethics guidelines has so far almost exclusively focused on conflicts within the scientific community. The statistics published by the ombudsmen show that the vast majority of cases concern conflicts between scientists concerning authorship or university appointments. The largest number of these cases concerned authorship and plagiarism (48/162 accepted cases), followed by cases concerning ownership of research equipment and of data (35/162) and those concerning obstruction of research (27/162).<sup>5</sup> The committee's dedication to conflicts among the scientific

---

4 According to the second report of the ombudsmen committee from June 2001, 58% of German universities had adopted such rules.

5 These statistics are from the first six years of the committee's work. They, as well as yearly reports can be found at [http://www1.uni-hamburg.de/dfg\\_ombud/](http://www1.uni-hamburg.de/dfg_ombud/). Given that researchers



community and the absence of cases concerning the rights of research subjects follows logically from the structure and procedures of the committee. Because the committee does not initiate investigations, it is naturally left with cases where colleagues accuse their peers and typically this concerns issues where one academic career is hindered by the other. Indeed, in Germany good practice appears to almost exclusively cover the rights of researchers and how they are treated by their community. Good scientific conduct thus becomes an ethos of scientific honesty towards one's colleagues rather than towards the research subjects. In short, it is unlikely that such voluntary rules that give priority to "self-monitoring" will be sufficient to promote research ethics for empirical research and teaching in the social sciences.

### **3. Social science research ethics in the United States and Britain**

#### *3.1 US: the legal approach*

In the US, federal regulations have made ethical standards for research involving human subjects mandatory since the early 1970s, in cases where research is conducted at federal institutions or is funded by federal agencies (National Research Act 1974).

The National Research Act, on which the current rules are based, was a reaction to abuses in human subjects research.<sup>6</sup> It led to the establishment of the *National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research*, which had two main tasks. First, it identified the basic ethical principles that should underlie the conduct of biomedical and behavioral research involving human subjects. Second, it developed guidelines assuring that such research was conducted in accordance with those principles. In 1978, the Commission established the "Ethical Principles and Guidelines for the Protection of Human Subjects of Research," better known as the "Belmont Report."

Important parts of the Belmont Report were included in the current legal framework for ethical research, the Code of Federal Regulations (CFR), in particular Title 45 CFR part 46 (The Code of Federal Regulations Governing

---

seem to appeal to the DFG ombudsmen rather to those of their own institutions (seen as too partial), the statistics published by the DFG are of some generality.

6 One of the most infamous cases of ethical misconduct of research is the Tuskegee Syphilis Study, which was a longitudinal project conducted between 1932 and 1972 by the US Public Health Service on poor, illiterate black men in rural Alabama. During this study 28 participants died.

the Protection of Human Subjects in Research). This framework was enacted in 1991 by the US Department of Health and Human Services (HHS), specifically its Office for Human Research Protections (OHRP). It includes the requirements for assuring compliance by research institutions,<sup>7</sup> for researchers obtaining and documenting informed consent, and for ethics review committees (Institutional Review Boards or IRBs) membership, function, operations, review of research, and record keeping.

In 1991, seventeen other Federal Agencies and Departments also adopted a uniform set of rules for the protection of human subjects, almost identical to Title 45 CFR part 46 (Subpart A).<sup>8</sup> This joint agreement on regulations is named the Federal Policy for Protection of Human Subjects, better known as the “Common Rule.”

The Common Rule is based on three fundamental principles for ethical research:

- Respect for a person’s autonomy: the researcher has to give adequate and comprehensive information about the research project and possible risks
- Beneficence: research has to maximize benefits for society and minimize risks for research subjects
- Justice: research must not exploit or ignore one group in order to benefit another group

Based on these principles, there are three core criteria for evaluating human subjects research: informed consent, risk-benefit assessment, and equitability of subject selection. Institutional Review Boards (IRBs) are the instrument for approving whether research is following these criteria. Most large universities and hospitals conducting research have established their own IRBs.

There are three different types of IRB (Parvizi et al 2007): (a) local IRBs are affiliated with the institution or organization conducting research, (b) central IRBs deal with large scale multi-site research, and (c) commercial IRBs are paid to review research with human subjects.<sup>9</sup> In the last few decades, the impact of IRBs on the research infrastructure has increased enormously. Indications of this development can be found in the increasing number of IRBs and the increase in their power. Although universities in the

---

7 Institutions normally make an agreement with the appropriate federal agency that funds their research. Most universities have an agreement with the HHS.

8 Further subparts of the 45 CFR part 46 outline rules for research on fetuses, neonates, and pregnant women (Subpart B); rules for research with prisoners (Subpart C); and rules regarding research involving children (Subpart D).

9 Commercial IRBs have become more common in recent years. The responsibilities of these IRBs as for-profit organizations are identical to those based at academic or medical institutions and they are governed by the same federal regulations.

US are generally confronted with numerous regulations and bureaucracies, the IRB system is the only one that has the direct power to stop, delay, or change the character of research (see Bledsoe et al. 2007).

At first glance it seems that the regulation of ethical research standards through the IRBs is an appropriate model, not only to ensure the protection of research subjects but also to bring binding ethical standards to the social sciences. However, the IRBs' practices for approving research projects in the social sciences is by no means undisputed. More to the point, IRBs have been directly criticized along three main lines.

First and most particularly, IRBs have been criticized as being inappropriate for the social science. Their composition and their requirements are seen to privilege research methods similar to the natural sciences. Indeed, the Common Rule regulations and the Belmont principles were developed with biomedical and laboratory science methods in mind. As Milne (2005) emphasizes, the type of research documentation to be brought to the IRB, such as informed consent protocols, asks for objectivity, prediction, and control rather than description, interpretation, and discovery. Using this approach, there would seem to be little room for qualitative forms of data collection and research. This general critique holds true despite some noteworthy exemptions from full IRB review in the case of research that is particularly relevant to social science investigation. Such exemptions apply, for instance, to research about educational practices or research involving the collection or study of existing data if publicly available or unidentifiable.<sup>10</sup>

Secondly, the research review boards have been criticized for their strong bent towards legal issues. As the process of research review focuses heavily on producing a legally valid written consent form, Bledsoe et al. (2007: 631) argue that the main goals of these reviews appears to be not so much to protect the research subjects but rather to deflect as much risk as possible from the institution. As a legal contract between the investigator and the university, the IRB protocol is an instrument to place as much legal responsibility on the investigator by defining as many risks as possible that have to be considered prior to research. In other words, universities turn to delegating legal risk to their faculty members.

Finally, ethics reviews have suffered from an externalization and professionalization of ethical problems from the point of view of the researchers. Faced with extensive IRB protocols, researchers tend to simply do their paperwork in the required manner, rather than thinking through the ethical issues related to their work.

---

10 For more detailed information, see NSF (National Science Foundation) "What exemptions of the Common Rule are most appropriate to social science research?" (<http://www.nsf.gov/bfa/dias/policy/hsfaqs.jsp>)

### 3.2 United Kingdom: The reformist approach

Since 2006, the explicit guidelines of the *Economic and Social Research Council* (ESRC), the main sponsor of social science research in the UK, have forced universities seeking its funding to consider ethical issues in research and teaching alike. These guidelines seek to establish rules suitable for the social sciences, stating that extant guidelines like those for medical research may not be appropriate for the social sciences with its diverse methodologies. They also seek to respond to new challenges in social sciences research ethics that arise from, among other things, interdisciplinary research, globalization, and technological change (see ESRC 2005).

This *Research Ethics Framework* (REF) is the result of consultations with the UK social science community, including other funding institutions and professional associations. The resulting six key principles of ethical research require (1) that research be designed, reviewed, and undertaken to ensure integrity and quality, (2) that research staff and subjects be fully informed about the purpose, methods, and intended possible uses of the research as well as its risks, (3) confidentiality of information and anonymity of respondents, (4) voluntary participation, (5) the avoidance of harm to research participants, and (6) independent research and explicit statements of conflict of interest or partiality. It is noteworthy that four of these key principles deal with the protection of research subjects (rather than with misconduct among peers).

The implementation of these ethical standards is delegated to universities or research institutes. Ensuring research ethics goes beyond a particular research project for which a research institution seeks funding. Indeed, only those institutions that have put mechanisms and procedures in place to ensure minimal ethical standards can apply for funding from the ESRC.

Although the ESRC does not impose a particular model to ensure ethical standards, it stipulates that the minimal mechanisms must include, most importantly, a REC that looks at ethical issues in research applications and monitors the implementation of the project.<sup>11</sup> Moreover, any application to the ESRC has to explain if and why it needs a review by the RECs. The reviewers from the ESRC have to comment on these ethical self-assessments in the proposal and may reject a proposal or give a conditional award only. Additional “incentives” to ensure ethical standards in social science research are provided by the possible loss of funding by the ESRC, even if other non-ESRC funded projects in a research institution breach ethical standards.

Among the minimal standards in the ESRC guidelines are, however, not only research ethics for the actual research process but also for training. At

---

11 Members of RECs need to be trained to deal with ethical issues and have to be compensated for their work. The REF leaves it open whether social science sub-RECs are to be created or if ethical issues in social science research are to be treated by the general RECs.

the very least, social science postgraduate training programs have to incorporate the range of issues addressed in the REF. This requires the development of minimum standards of training and competence in ethical issues over time. According to the REF, such minimum training requirements are likely to include: training for individual researchers, training for members of local and institution-wide RECs, and training for postgraduate students in local ethics review requirements – in addition to more general ethics training as well as training for undergraduate students whose projects may require an ethics review (see ESRC 2005, 16).

As a relatively recent framework compared to the US system, the REF is probably more suitable to serve as a model for research ethics in the social sciences for four main reasons. First, the ensuring of research ethics is delegated to research institutions (although it is monitored by the ESRC). This decentralized approach could be more suitable for Germany because it would respect the independence of universities. Second, because the REF seeks to decrease delays and unnecessary efforts, the evaluation and approval of the REC is not necessary for the actual application, but only for the beginning of the project. Third, the REF not only creates negative incentives but introduces ethical issues into training. The purpose thus appears to be not to simply create a lengthy procedure to be complied with on the way to obtaining funding, but also to contribute to a research culture where ethical issues are viewed as an important part of research and training. Fourth, the REF recognizes explicitly that qualitative methodology may require a different type of ethical review than quantitative methodology.<sup>12</sup>

#### 4. Discussion

The three ethics frameworks for the social sciences discussed here vary widely in their treatment of both of our two key criteria – the protection of research subjects and their appropriateness for the social sciences. They also differ in how they address various important sub-issues such as the degree of bindingness, the locus of implementation, and the weight given to raising awareness of ethical issues in the training of researchers.

Obviously, the German approach is the most underdeveloped one. Ethical principles are strongly considered in (bio)medical research, but this has not been extended to the social sciences. Those guidelines that address the whole scientific community in Germany, such as the DFG guidelines on

---

12 On the negative side, obviously, are the greater costs for research institutions, since they need to create new bodies – the RECs – and compensate its members for their work. The ESRC, however, argues that the cost of reviewing ESRC-funded projects is also eligible for funding.

good scientific practice, focus almost exclusively on ethical behavior among peers rather than on the protection of research subjects. Providing few constraints, they have moreover received little promotion within the university. The ethics codes of professional associations do include sections on the protection of research subjects. No ethics reviews of research projects protect, however, the rights of research subjects *ex ante* and their ethics committees as well as the DFG ombudsmen are, structurally, unlikely to be summoned by research subjects *ex post*.

The US approach, in contrast, with its requirement for projects to have approval by IRBs, is highly protective of research subjects in the natural and social sciences alike. At the same time, the framework does not consider methods specific to the social sciences that make the issue of informed consent particularly complicated in its implementation. As a result of the origin of the IRBs in and membership bent towards the natural sciences, US social science research tends to lean towards “standard methods” in order to receive IRB approval. In this way, research ethics has a strong and not always beneficial effect on the content of social science research. Moreover, the significant bureaucratic work involved in getting IRB approval makes data gathering cumbersome and is therefore only encouraged at the post-graduate level.

The UK approach seeks to strike a balance. Given that the largest social science funding institution makes ethics reviews and ethics committees a requirement, it gives considerably more protection to the research subjects than the German system. Being designed for social science research, it is also much more open towards qualitative methodology than the US approach. Requiring an ethics review only for approved research projects, it also entails a less lengthy procedure than the US model, even though it requires researchers to think about ethical issues (i.e., the type of necessary review) when designing their project. Of the three approaches, it is also the one that most energetically stresses the need to raise the awareness of ethical principles during training.

## 5. Recommendations

Research ethics is about social responsibility and thus goes beyond a simple set of legal regulations. An ethics framework should thus give priority to raising awareness of ethical principles in research. This means that research ethics, and, more importantly, learning to think about the ethical dimension of their work, should be an integral part of the training of young researchers. Germany would benefit from a new research ethics framework for the social sciences. This framework should focus on protecting the rights of the re-

search subjects and encompass data access, gathering, and processing. But it should also make it a priority to accommodate social science methodologies. Like the UK approach, it could be enforced by making it a mandatory step in the funding process. The US example shows that legal requirements may create many bureaucratic hurdles for research as well as having an undesired streamlining effect on its content.

This ethics framework could be modeled on the UK, but should be developed in consultation with the relevant professional associations, key funding institutions, universities, and independent research institutions in Germany. It should be reviewed upon request, following methodological innovations.

A German research ethics framework should give the responsibility for implementation to the universities. Independent research institutes should cooperate with the universities. At present, social science departments have neglected questions relating to research ethics in training and research practice. “Local” ethics committees with alternating members would bring the discussion and consideration of ethical principles into the universities. Such a system would integrate researchers into the implementation process of ethical standards rather than suspecting them *a priori* of misconduct.

## References:

- Bledsoe, C.H./Sherin, B./Galinsky, A.G./Headley, N.M./Heimer, C.A./Kjeldgaard, E./Lindgren, J./Miller, J.D./Roloff, M.E. and Uttal, D.H (2007): Regulating Creativity: Research and Survival in the IRB Iron Cage. *Northwestern University Law Review* 101 (2), 593-641.
- Carpenter, D. (2007): Institutional Review Boards, Regulatory Incentives, and Some Modest Proposals for Reform (2007). *Northwestern University Law Review* 101 (2). *Minnesota Legal Studies Research Paper No. 06-64*. <http://ssrn.com/abstract=948739>. [Last visited: 01/18/2010].
- Department of Health and Human Services – Office for Protection from Research Risks, Code of Federal Regulations Title 45, Public Welfare, Part 46, Protection of Human Subjects, Revised: June 18, 1991, Effective: August 19, 1991. <https://irb.llnl.gov/appendices/Appendix03.pdf>. [Last visited: 01/18/2010].
- DFG (1997): *Vorschläge zur Sicherung guter wissenschaftlicher Praxis*. Bonn.
- ESRC (2005): *Research Ethics Framework*. [http://www.esrcsocietytoday.ac.uk/ESRCInfoCentre/Images/ESRC\\_Re\\_Ethics\\_Frame\\_tcm6-11291.pdf](http://www.esrcsocietytoday.ac.uk/ESRCInfoCentre/Images/ESRC_Re_Ethics_Frame_tcm6-11291.pdf). [Last visited: 01/18/2010].
- Greely, H.T. (2008): *Collecting Biomeasures in the PSID: Ethical and Legal Concerns*. Mimeo.
- Haggerty, K.D. (2004): Ethics Creep: Governing Social Science Research in the Name of Ethics. *Qualitative Sociology* 27 (4), 391-414.
- Lindgren, J./Murashko, D. and Ford M.R. (2007): Symposium, Foreword: Symposium on Censorship and Institutional Review Boards. *Northwestern University Law Review* 101 (2), 399-403
- Milne, C. (2005): Overseeing Research: Ethics and the Institutional Review Board, *Forum Qualitative Sozialforschung/ Forum: Qualitative Social Research* 6 (1), Art. 41. <http://nbn-resolving.de/urn:nbn:de:0114-fqs0501412>. [Last visited: 01/18/2010].
- National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, The Belmont Report, Ethical Principles and Guidelines for the Protection of Human Subjects of Research, April 18, 1979. <http://ohsr.od.nih.gov/guidelines/belmont.html>. [Last visited: 01/18/2010].
- National Science Foundation (NSF): *Frequently Asked Questions and Vignettes, Interpreting the Common Rule for the Protection of Human Subjects for Behavioral and Social Science Research*. <http://www.nsf.gov/bfa/dias/policy/hsfaqs.jsp>. [Last visited 01/18/2010].
- Parvizi, J./Tarity, T./Conner, K. and Smith, B. (2007): Institutional review board approval: Why it matters. *The Journal of Bone and Joint Surgery*, 89-A (2), 418-426.
- United States Department of Health & Human Services/Office for Human Research Protection (OHRP): *Frequently Asked Questions*. <http://www.hhs.gov/ohrp/45CFRpt46faq.html>. [Last visited 01/18/2010].
- Schaar, P. (2010): Data protection and statistics – a dynamic and tension-filled relationship. [In this publication].
- Weber, K. (2010): New Methods in Human Subjects Research – Do We Need a New Ethics? [In this publication].



#### ***IV. FIELDS***

##### **1. MIGRATION AND DEMOGRAPHY**



## 1.1 Migration and Globalization

*Martin Kahanec and Klaus F. Zimmermann\**

---

\* The views expressed are the authors' alone and do not necessarily correspond to those of the German Data Forum (RatSWD). We thank Uwe G. Rehfeld and the participants to the Second Workshop of the German Data Forum (RatSWD) for helpful comments.

Contact:

Klaus F. Zimmermann  
IZA, P.O. Box 7240  
53072 Bonn  
Germany  
e-mail: Zimmermann[at]iza.org

## Abstract

The international migration of people is a momentous and complex phenomenon. Research on its causes and consequences requires sufficient data. While some datasets are available, the nature of migration makes it complicated to use them for scientific research. There is currently no existing dataset that effectively captures international migration trajectories. To alleviate these difficulties, we recommend: (1) the international coordination of data collection methodologies and standardization of immigrant identifiers; (2) a longitudinal approach to data collection; (3) the inclusion of adequate information about relevant characteristics of migrants in surveys, including retrospective information; (4) minimal anonymization; (5) immigrant boosters in existing surveys; (6) the use of modern technologies and facilitation of Data Service Centers; and (7) making data access a priority of data collection.

Keywords: migration, immigrants, data collection, data access, data infrastructure

## 1. Introduction

The international migration of people lies at the core of the ongoing process of globalization. People migrate to improve their economic prospects, ensure a more secure living environment, reunite with their family members, or avoid persecution in their country of origin. For these and other reasons 3 percent of the world's population found themselves on an international migration trajectory in 2005. Since a large proportion of these migrants head towards developed countries, the share of international migrants in these countries reached as much as 9.5 percent in 2005.<sup>1</sup> These individuals experience not only important economic and social consequences of their move, but also psychological ones. Migration may involve a new job with higher pay, the loss of old social ties and the establishment of new ones, as well as the psychological costs of missing the homeland.

Migration, however, does not only affect the fates of those who are directly involved. Various effects emerge at the interface of migrant and native populations. Immigrants may bring with them new cultures or preferences, they may compete for certain jobs and create others, or claim publicly financed social security benefits. More broadly, migrants contribute to a more efficient allocation of resources and often become a driving force behind knowledge transfer and technological advancement. All these effects have repercussions for the native population, who may react to migrant inflows not only in determining their current actions, but also in making

---

1 See United Nations, Department of Economic and Social Affairs, World Migrant Stock: The 2005 Revision Population Database.

long-term investment plans, such as those concerning education. Finally, natives may view immigrants positively or negatively and shape their attitudes accordingly.

Migration is a dynamic phenomenon involving many twists and turns. Driven by a multitude of possible reasons, migrants may move temporarily or permanently, transnationally and nationally, individually or in groups, return to their countries of origin or migrate to another country, or move between two or more countries in a circular way. The complex underlying processes driving migration and its effects have attracted significant and growing attention from scientists. Chiswick (1978) and Borjas (1985) have pioneered the scientific work on immigrant adjustment in host societies, highlighting the significance of the host country experience and stressing the importance of cohort effects, country of origin, religion, education, as well as a number of demographic characteristics such as age and gender. From a different perspective, research on the migration decision has been inspired and advanced by Harris and Todaro (1970), Becker (1964), Mincer (1978), and Borjas (1985). Immigrant self-selection discussed by Borjas (1987) and Chiswick (1999) implies the need for specific techniques (Heckman 1979) to consistently evaluate causal mechanisms behind immigrant adjustment.

The impact of immigration on the host labor market has been modeled by Chiswick, Chiswick, and Karras (1992) and Chiswick (1998). A large body of empirical literature, summarized by Kahanec and Zimmermann (2009), provides mixed evidence on the sign and determinants of these effects on wages and employment.<sup>2</sup> More recently, the roles of intermarriage (Meng and Gregory 2005), citizenship (Bratsberg et al. 2002), social networks (Munshi 2003), and attitudes (Bauer et al. 2000; Kahanec and Tosun 2009) pertaining to immigrant adjustment have received significant attention. The concept of ethnic identity has been extended by Constant and Zimmermann (2008), who elaborate on how attachment to the country of origin and the host country affect immigrant adjustment.

Although measuring the effects of migration is by no means an easy or straight-forward job, migration is a phenomenon that undoubtedly affects the well-being of the whole society and as such, has become an important and sensitive policy issue. Questions about the labor market consequences of migration, immigrant adjustment in host societies, and welfare competition have received particularly significant policy attention.

Understanding the causes and effects of international migration flows requires a sound and in-depth analysis. The need for such analysis is most conspicuous in the study of causal relationships, as these are difficult to establish empirically and their misrepresentation compromises both scientific and policy analyses. In fact, it may lead to incorrect policy recommendations,

---

2 The evidence on migration effects in source countries is mainly related to remittances (e.g., Barham and Boucher 1998), and wage and employment effects (Brücker 2007).

which may in turn lead to unpredictable consequences or even effects contrary to those intended. Since such analysis is impossible without high-quality data, such data are indispensable for policy analysts as well as scientists.

## 2. Relevant and available data

Despite the general scarcity of migration data, scientists and analysts have been able to use some existing survey or administrative datasets as well as small-scale dedicated survey data to study migration issues. While these datasets have facilitated valuable research, missing variables, excessive anonymization, and flaws in data collection design often compromise scientists' efforts to broaden and deepen our knowledge of migration causes and effects. In this section we focus on some large-scale datasets collected at the European level, since they, in contrast to small-scale surveys, have an intrinsic potential to provide the necessary transnational, longitudinal, and systematic data collection framework.

There are four extensive datasets that in some dimension provide coverage of European migration trajectories: the European Community Household Panel (ECHP), the EU Statistics on Income and Living Conditions (EU-SILC), the EU Labour Force Survey (EU-LFS), and the OECD/SOPEMI (*Système d'Observation Permanente des Migrations*) dataset. Each of these datasets contains information about demography, labor force participation, employment, unemployment, self-employment, and educational attainment of immigrants. In addition, the European Social Survey (ESS) covers people's attitudes toward immigrants as well as their voting preferences, thus addressing migration indirectly.

Table 1 depicts the character of these datasets, highlighting some of their strengths and weaknesses. We can identify at least three major gaps in the available data.<sup>3</sup> First, these datasets provide either none or only a very limited account of migration trajectories. Transnational migration trajectories may involve simple or repetitive moves between two or more countries with temporary periods of residence of varying length as well as permanent moves. It is almost impossible to track such trajectories – with all their spells, stops, and circularities – within Europe or between Europe and third countries. In particular, little or no information is available on migrant experience prior to their arrival in the country of current residence or their intentions for further moves. Secondly, the data typically permit determining immigrant status based on an individual's citizenship and country of origin,

---

3 See also the discussion in Bauer and Zimmermann (1998).

neglecting the large groups of people with an immigrant background who are native citizens or those with dual citizenships. Finally, anonymization often renders any valuable analysis impossible, for example when immigrants from very different origins (e.g., Zimbabwe and Japan) are grouped into one category (e.g., non-EU).

### 3. Data access issues and needs

Inadequate access to existing datasets is one of the most limiting factors for scientific and policy analysis. Due to restrictive data access policies, a lack of interest on the side of the officials responsible, misinterpreted data protection rules, or simply the lack of adequate data access infrastructure, the use of datasets for scientific and policy purposes is, in general, severely limited. Since migration is, by definition, a transnational and dynamic phenomenon (i.e., involving one-way as well as repeat, sequential, and circular movement between more countries), its proper analysis requires a combination of information from multiple countries and across multiple periods.<sup>4</sup> Therefore, restrictions on data access and a lack of coordination in establishing access rules are particularly detrimental to the analysis of migration issues. Below we list some of the most pressing issues that obstruct availability of data for migration analysis and determine the needs concerning collection of adequate data on migration.

One of the main problems is that identifying and defining migrants in the existing datasets is not a trivial matter. The migration background, a foreign origin (foreign birthplace), citizenship, or ethnicity can be used to determine whether or not someone is an immigrant. Unfortunately, only a subset of this information, if any, is available in existing datasets. Only rarely can one identify first, second, and further generations of immigrants, citizens and non-citizens, and distinguish immigrants of different origin and ethnicity.

It is even more seldom that it is possible to obtain information that can be used to characterize migration trajectories. Perhaps with the exception of length of stay in the host country, pre-migration experience, tracking of all migration moves, or migration trajectories of family members (spouses) are rarely available. While the lack of data describing migration trajectories of those who make more frequent, possibly circular, moves is a general problem, it is particularly problematic in the case of highly skilled migrants, since these are the most fluid and mobile segment of the migrant population.

---

4 Bauer, Pereira, Vogler and Zimmermann (2002) have merged Portuguese data and German data on Portuguese migrants to be able to compare migrants in the sending and a receiving country. See Crul and Vermeulen (2006) for another project in this spirit.



Other relevant and often missing information include language, religion, and attachment to the host society and the country of origin.

A further and related problem is that the effects of out-migration are hard to capture, since we typically do not observe people who leave or record their characteristics (they do not de-register and are in a different country when data are collected). In fact, this deficiency creates problems for the analysis of the entire population as well, since it compromises the representativeness of datasets. For example, according to the Weekly Report of the German Institute for Economic Research (“Wochenbericht des DIW”) (2008, 382), doubts have arisen in Germany as to whether the official census statistics still represented the actual reality of the German population. Since the German national census data has only been based on registers since 1987 – which depend on the proper *registration* and *deregistration* of people – those who leave the country and do not deregister are erroneously counted. An example of the magnitude of the measurement error which can result from failing to track out-migration of those who have not deregistered was revealed in a clean-up of the data from the German Central Register of Foreigners (*Ausländerzentralregister*) in 2004, which showed that the official census statistics had overstated the number of foreigners in Germany by about 600,000.

Another problem with the current situation is that most datasets are representative of the total population and contain a limited number of observations. While this is not necessarily a problem in other contexts, in the context of migration it often implies insufficient samples of the immigrant population. In addition, many datasets are cross-sectional and thus do not capture the dynamic nature of migration. In particular, the snapshot picture that such datasets provide can only capture the most recent move and cannot distinguish some important effects, such as those of host country experience and immigrant cohort on immigrant adjustment.

Finally, knowledge of migration intentions and reasons, and their relationship to actual migration decisions is indispensable for predicting future migration flows as well as for understanding the social and economic outcomes of migrants in the host societies. Precise estimates of the directions and characteristics of such flows are crucial for designing effective and efficient immigration policies, for instance in the context of EU enlargement. The intention to stay, namely, whether migrants perceive their situation as temporary or come to settle in the host country permanently, carries important consequences for their labor market behavior and thus the effects they have on the host economy. Similarly, migrants who come for economic reasons and those who come as refugees or asylum seekers have very different labor market opportunities as well as intentions in the host country.

These issues concerning the availability of and access to adequate migration data define the primary needs concerning the collection of relevant

data. In particular, collected data should properly identify migrants and people with immigrant backgrounds, and contain sufficient samples of migrants. They should cover (transnational) migration trajectories and, in particular, capture pre-migration experience and out-migration, as well as measure intentions and reasons for migration.

#### **4. Future developments and challenges**

The enlargement of the European Union and the concurrent expansion of the European Economic Area as well as the persistent economic and social hardship and insecurity in large parts of the world will continue to fuel substantial international flows of people. High-quality data are and will remain a key ingredient to understanding the causes and effects of these migration flows. Given the traditional prominence of quantitative techniques in economics and the growing emphasis on such techniques in other social sciences, especially sociology, we can project increasing demand for such data among scientists in the future. This demand will be further intensified by the increasing need for well-founded policy analysis at the European and national levels. Another contributing factor may be the business sector, which may seek to exploit the potential benefits from precise information about their current and potential customers.

The provision of high-quality migration data is in general insufficient, although it has somewhat improved over the last decade or two. This improvement has been enabled by the emergence of advanced information and data management technologies that can facilitate a wide access to existing datasets. This development particularly relates to a group of international institutions that have started to provide online access to some of their datasets (European Union, World Bank, ILO,<sup>5</sup> UN) as well as private and non-governmental organizations (IZA)<sup>6</sup> that use innovative technologies to promote access to own and third-party datasets. While some improvements have been made at the national level, government institutions still lag behind in data access provision. More recently, some remarkable developments have taken place, however, involving a partnership between public and non-governmental or private institutions aiming at a wider dissemination of valuable data collected by public institutions. For instance, the International Data Service Center of the IZA offers onsite computing via ultra thin access, and remote computing by means of a remote computing solution (JoSuA)

---

<sup>5</sup> International Labour Organization.

<sup>6</sup> Institute for the Study of Labor (IZA, *Forschungsinstitut zur Zukunft der Arbeit*).

facilitating the use of scientific use data of the German Federal Statistical Office.<sup>7</sup>

These positive developments should not hide the difficult reality of migration research and analysis as concerns data availability. Besides the various difficulties that migration researchers face regarding identification of migrants in existing datasets as well as a lack of relevant information about them, virtually no existing dataset has the necessary transnational and longitudinal perspective to capture complete migration trajectories. Thus, the key challenge in this respect is to track migrants and their migration experience as they move internationally. The associated practical challenge is to coordinate data collection methodologies across Europe and, even more complicated, between Europe and third countries.

## 5. Conclusions and recommendations

This essay summarizes some of the key problems and challenges related to the availability of data for the study of migration issues. Having considered the long-standing as well as more recent developments in migration research, it is clear that access to data of good quality, harmonized across time and countries, is one of the main bottlenecks hindering advances in our understanding of the causes and effects of migration. To alleviate this problem, there are a number of policy tools that may help.

First, coordination of data collection methodologies and standardization of immigrant identifiers across the EU would facilitate international comparability. It is necessary to harmonize data collection methods so that migration trajectories in Europe-wide datasets can be observed. In particular, unique individual identifiers need to be traceable across European countries. An open method of coordination, transparent indicators, benchmarking, and an efficient exchange of best practices seem to be the way to go in this regard. This also involves merging datasets transnationally and across time, including proper harmonization and linking of data, records, and topics. In particular, given the advancement in data management and storage techno-

---

7 The International Data Service Center of the IZA, one of the Data Service Centers facilitated by the “KVI Commission,” offers an integrated service which consists of a metadata portal and a remote computing solution. The International Data Service Center of the IZA’s metadata service comprises a detailed, in depth, searchable and standardized information and documentation service on a growing number of datasets currently in the areas of employment and wages, education and training, and demographics and migration. The International Data Service Center of the IZA remote computing solution, known as JoSuA, facilitates usage of restricted datasets bridging the otherwise wide gap between legal constraints and scientific freedom without violating the former or constraining the latter. For further details see Schneider and Wolf (2008).

logies, this objective entails not only prospective but also retrospective harmonization and merging of datasets as well, involving digitalization of old datasets whenever necessary.

Second, whenever possible, a longitudinal approach should be adopted, both to facilitate the separation of spurious effects driven by unobserved cross-sectional variation from true causal relationships as well as to capture the dynamic nature of migration. In this regard, one should also consider extending selected existing cross-sectional datasets by surveying the covered individuals in one or more additional waves.

Third, adequate information about the relevant characteristics of migrants – experience in the host society (years since migration), country of origin, citizenship, ethnicity, language, religion, attachment to the host society and the country of origin, and migration intentions and reasons – is requisite. For example, it is of key importance to distinguish temporary and permanent migrants as well as economic migrants from those that come as refugees or asylum seekers, or as tied movers. Retrospective questions in survey questionnaires are necessary to track migrants' pre-migration experience (i.e., experience prior to the last observed move).

Fourth, anonymization should be limited to the smallest possible degree. As an option, alternative anonymization procedures could be applied to the same dataset, allowing two or more versions to be accessible to the researcher, each facilitating research on different research questions.

Fifth, immigrant boosters in existing surveys with a well-defined control group would facilitate sufficient immigrant sample sizes.

Sixth, online Data Service Centers, data registers and metadatabases can provide an invaluable service to the research community. In fact, the Internet is itself becoming a rich source of data and a tool to collect new data that still needs to be properly exploited.

Seventh, the use of modern data information technologies should be promoted to facilitate the collection, management, and storage of good quality data, as well as, importantly, enabling access to it. As part of this objective, the creation of Data Service Centers facilitating prudent access to such data is desirable.

Finally, the facilitation of data access for researchers should be embraced as one of the main objectives of data collection. Adequate effort by all the actors involved is necessary not only to facilitate knowledge about migration as such, but also, to the extent that suitable policies are adopted, to improve the welfare of substantial numbers of people who are directly or indirectly affected by migration.

Table 1: Datasets

Dataset	Type	Years	Countries	Measures of immigrant status	Weakness (selected)
ECHP	Longitudinal	1994 - 2001	EU 15	Year of arrival (region/country), country of birth, first and second citizenship, mother tongue	No information on immigrant experience prior to his/her arrival to the country of present residence.
EU-SILC	Longitudinal	2004 - 2006	Until 2004 EU15 Since 2006 EU25	Country of birth, citizenship (first)	Anonymization leading to a mixing of immigrants from very different origins.
EU-LFS	Survey	1983 - 2006	BE, CZ, DK, DE, EE, GR, ES, FR, IE, IT, CY, LV, LT, LU, HU, MT, NL, AT, PL, PT, SI, SK, FI, SE, UK, BG, RO (+HR, TR, IS, NO, CH) OECD	Nationality (citizenship), years of residence, country of birth (anonymized), country of residence one year before the survey	Anonymization leading to a mixing of immigrants from very different origins.
OECD/SOPEMI	Macro-data	1983-2008	OECD	Stocks of foreign nationality and foreign born populations, country of birth, flows of foreign-born workers	While the dataset provides aggregate data, no information about the individual characteristics of migrants is available.
ESS	Cross-sectional	2002, 2004, 2006	EU25	Voting preferences, attitudes toward immigrants and ethnic minorities	The cross-sectional nature of the dataset does not capture the dynamic nature of migration.

## References:

- Barham, B. and Boucher, S. (1998): Migration, Remittances and Inequality: Estimating the Net Effects of Migration on Income Distribution. In: *Journal of Development Economics* 55 (2), 307-331.
- Bauer, T./Lofstrom, M. and Zimmermann, K.F. (2000): Immigration Policy, Assimilation of Immigrants, and Natives' Sentiments towards Immigrants: Evidence from 12 OECD Countries. *Swedish Economic Policy Review* 7 (2), 11-53.
- Bauer, T. and Zimmermann, K.F. (1998): Causes of International Migration: A Survey. In: Gorter, C./Nijkamp, P. and Poot, J. (Eds.): *Crossing Borders: Regional and Urban Perspectives on International Migration*. Aldershot.
- Bauer, T./Pereira, P.T./Vogler, M. and Zimmermann, K.F. (2002): Portuguese Migrants in the German Labor Market: Performance and Self-Selection. *International Migration Review* 36 (2), 467-491.
- Becker, G.S. (1964): *Human Capital. A theoretical and empirical Analysis*. New York.
- Borjas, G.J. (1985): Assimilation, Changes in Cohort Quality, and the Earnings of Immigrants. *Journal of Labor Economics* 3 (4), 463-489.
- Borjas, G.J. (1987): Self-Selection and the Earnings of Immigrants. *American Economic Review* 77 (4), 531-553.
- Bratsberg, B./Ragan, J.F. and Nasir, Z.M. (2002): The Effect of Naturalization on Wage Growth: A Panel Study of Young Male Immigrants. *Journal of Labor Economics* 20 (3), 568-597.
- Brücker, H. (2007): Migration after the EU's Eastern Enlargement: Who Wins, Who Loses? [Paper presented at Second IZA Migration Workshop: EU Enlargement and the Labor Markets, Bonn 7.-8. September 2007].
- Chiswick, B.R. (1978): The Effect of Americanization on the Earnings of Foreign-born Men. *Journal of Political Economy* 86 (5), 897-921.
- Chiswick, B.R. (1998): The Economic Consequences of Immigration: Application to the United States and Japan. In: Weiner, M. and Hanami, T. (Eds.): *Temporary Workers or Future Citizens? Japanese and U.S. Migration Policies*. New York.
- Chiswick, B.R. (1999): Are Immigrants Favorably Self-Selected? *American Economic Review* 89 (2), 181-185.
- Chiswick, C.U./Chiswick, B.R. and Karras, G. (1992): The Impact of Immigrants on the Macroeconomy. *Carnegie-Rochester Conference Series on Public Policy* 37 (1), 279-316.
- Constant, A. and Zimmermann, K.F. (2008): Measuring Ethnic Identity and Its Impact on Economic Behavior. *Journal of the European Economic Association* 6 (2-3), 424-433.
- Crul, M. and Vermeulen, H. (2006): Immigration, education and the Turkish second generation in five European nations: a comparative study. In: Parsons, C.A. and Smeeding, T.M. (Eds.): *Immigration and the Transformation of Europe*. Cambridge.
- Harris, J.R. and Todaro, M.P. (1970): Migration, Unemployment and Development: A Two-Sector Analysis. *The American Economic Review* 60 (1), 126-142.

- Heckman, J.J. (1979): Sample Selection Bias as a Specification Error. *Econometrica* 47 (1), 153-161.
- Kahanec, M. and Tosun, M. (2009): Political Economy of Immigration in Germany: Attitudes and Citizenship Aspirations. *International Migration Review* 43 (2), 263-291.
- Kahanec, M. and Zimmermann, K.F. (2009): International Migration, Ethnicity and Economic Inequality. In: Salverda, W./Nolan, B. and Smeeding, T.M. (Eds.): *Oxford Handbook on Economic Inequality*. Oxford.
- Meng, X. and Gregory, R.G. (2005): Intermarriage and economic assimilation of immigrants. *Journal of Labor Economics* 23 (1), 135-175.
- Mincer, J. (1978): Family Migration Decisions. *Journal of Political Economy* 86 (5), 749-773.
- Munshi, K. (2003): Networks in the Modern Economy: Mexican Migrants in the U.S. Labor Market. *The Quarterly Journal of Economics* 118 (2), 549-599.
- Schneider, H. and Wolf, C. (2008): Die Datenservicezentren als Teil der informationellen Infrastruktur. In: Rolf, G./Zwick, M. and Wagner, G.G. (Eds.): *Fortschritte der informationellen Infrastruktur in Deutschland. Festschrift für Johann Hahlen zum 65. Geburtstag und Hans Jürgen Krupp zum 75. Geburtstag*. Baden-Baden.
- Weekly Report of the German Institute for Economic Research (Wochenbericht des DIW) Nr. 27-28/2008, DIW Berlin.





## 1.2 Migration and Statistics

*Sonja Haug*

Contact:

Sonja Haug  
Hochschule Regensburg  
Fakultät Angewandte Sozialwissenschaften  
Postfach 12 03 27  
93025 Regensburg  
Germany  
e-mail: [sonja.haug\[at\]hs-regensburg.de](mailto:sonja.haug[at]hs-regensburg.de)

## Abstract

The field of empirical research on migration and integration encompasses a wide range of research questions, theoretical approaches, and datasets. Research based on official statistics has to deal with diverse datasets for information on migration and foreign populations, resulting in miscellaneous statements. Recent developments in official statistics have concentrated on the improvement of data quality. The 2011 Census and the projected creation of a central population register are both important issues, for example, for the sampling and weighting of migrants in surveys. The concept of migration background, too, as it has become integrated into the German Microcensus, represents a major development in population statistics and is now widely accepted. This report recommends implementing questions on migration background into the 2011 Census. The most important and accessible datasets in the field of empirical integration research are the German Microcensus and the German Socio-Economic Panel (SOEP, *Sozio-oekonomisches Panel*); there is still untapped potential for analysis. The addition of a migrant sample as a supplement to large surveys is another valuable innovation. The most important challenges for empirical research in migration and integration are the development of sampling methods for migrant population (including onomastics and topomastics), studies of new and small migrant groups, research projects focused on the country of origin, longitudinal migrant surveys, and the development of adequate tools for measurement.

Keywords: population, migration, integration, migration background

## 1. Research questions

The field of empirical migration and integration research is characterized by a wide range of research questions, theoretical approaches, and datasets.

### *1.1 Established research questions*

Migration research tends to center around a group of well-established questions: how many migrants are on the move? What are their countries of origin and destination? What are the determinants of migration? Integration research in Germany has paid particular attention to the population of “guest workers.” Empirical research has focused on a descriptive analysis of integration (Mehrländer et al. 1996), or on explaining the migration and integration process (Esser 1980).

### *1.2 Theoretical developments and new research questions*

Migration research today is more differentiated and has multiple points of focus. New research fields in international migration include: migration from Central and Eastern European countries, questions relating to demography and the concept of “replacement migration,” return migration, and irregular migration. New interdisciplinary approaches in sociology or economics employ theories of social capital and social networks to explain migration decisions (e.g., Boswell and Mueser 2008). Transnational migration, chain migration, circular migration, and migrant communities are other important areas of research (Pries 1997). Theoretical and empirical progress in this area allows, for example, for an analysis of international migration dynamics from a development perspective using a micro-level decision model and event history methodology (Massey et al. 2008). New developments in migration and integration theory have also arisen through discussions about a general model for intergenerational integration (Esser 2008; Kalter 2008). Current research concentrates on the different dimensions of integration and the mechanisms underlying the observable types of integration (e.g., Kalter 2008b).

Empirical research on migrant integration is an emerging field and frequently overlaps with issues like social inequality and exclusion (Kalter 2003). Theories relating to the sociology and economy of migration are increasingly being incorporated into this area of research. In addition, the integration of the second generation of migrant families is a core issue (Haug and Diehl 2005). The Programme for International Student Assessment (PISA) study and other research studies highlight the problems faced by migrant children in the educational system (Stanat and Christensen 2006; Kristen 2005); labor market integration is yet another topic of inquiry (Granato 2003). New research can be said to focus on the transition from primary to secondary school and/or the transition from vocational education to occupation. Analysis of the integration process is a key trend in this area, and panel surveys and longitudinal studies play an ever greater role in this, although adequate datasets are rare. Evidence shows that naturalized migrants integrate more successfully than non-naturalized migrants, but they still differ from people without a migration background (Haug 2002; Salentin and Wilkening 2003). As a result of these insights, the formerly widespread concept of “foreigners” is losing ground, and the concept of “migration background” becoming generally accepted (for an overview of measurement aspects, see Diefenbach and Weiß 2006).

Research on internal migration is another topic to emerge in recent years, for example, looking at migration decision making in terms of the labor market and the life cycle (Huinink and Kley 2008; Kalter 1997; Wagner 1989; Windzio 2004).

### *1.3 Policy-related research questions*

The analysis of legal migration and the estimation of illegal migration is one of the central policy-related issues in migration research (BAMF 2010). In Germany, the national contact point of the European Migration Network (EMN) is the Federal Office for Migration and Refugees (BAMF, *Bundesamt für Migration und Flüchtlinge*). Their aim is to improve the availability of and access to information on migration and asylum at the EU level; to facilitate work on annual policy reports, small scale studies, and annual reports on migration, asylum, and return migration statistics; and to support the policy and decision-making process. The target population for policy-related surveys is made up of former labor migrants and their families (Venema and Grimm 2002; Babka von Gostomski 2008; Weidacher 2000) or ethnic German repatriates (Haug and Sauer 2007), while special research projects have a narrower focus, such as studies on female migrants (Boos-Nünning and Karakasoglu 2006), the “second generation” (Haug and Diehl 2005), or Muslims (Brettfeld and Wetzels 2007). Integration policy-makers often tend to think of migration in terms of deficits, such as deficits in German language knowledge, in educational success, or in the labor market integration of migrants. What is needed, however, is a totally new approach towards prioritizing resources for migrants in areas such as language and professional skills. One emerging field of policy-related data analysis is the development of integration indicator sets in municipalities, the federal states, and on the national level (Siegert 2006; Worbs and Friedrich 2008; KGSt 2006; Filsinger 2008). The BAMF Integration Report Working Papers series provides an overview of official statistics and empirical social research covering a wide range of fields of integration. The Federal Government Commissioner for Migration, Refugees, and Integration has published a set of integration indicators and will publish a more detailed report at a later date. Comprehensive data analysis on integration aspects is also conducted in different fields of policy-related research; examples include the official reports on education (*Bildung in Deutschland*), poverty (*Armuts- und Reichtumsbericht*) or families (*Familienbericht*).

## 2. Status quo: Databases and access

Different databases are available for different research purposes (Diehl and Haug 2003; Haug 2005; Schönwälder et al. 2008).

### 2.1 *Databases for research into immigration and emigration (population flow)*

Since registration with the local authorities is obligatory in Germany, legal migration can be measured directly. Flow data result from decentralized local resident registration offices, which supply aggregate data on immigration and emigration – categorized by citizenship, country of origin or destination, age, and gender – to the Statistical Offices of the *Länder* and to the Federal Statistical Office. The lack of reliable figures on the number of irregular migrants leads to an underestimation of immigration and the size of the foreign population. Moreover, register information generally produces case-based statistics, so official aggregate migration statistics tend to overestimate migration. The restructuring of migration statistics for foreigners is ongoing, with the result that persons crossing the border several times a year are no longer counted multiple times. The migration data does not include specifics on the duration of stay or the residence permit status, so it is not possible to differentiate between long-term and short-term migrants.

A second source of migration flow data is the Central Register of Foreigners (AZR, *Ausländerzentralregister*; BAMF 2010). The major drawback of this register is that it only records the immigration of foreigners, and it cannot therefore be considered a comprehensive migration register. The main advantage of the register is that it provides person-based statistics and the option of distinguishing between short-term and long-term migrants (persons who have stayed for at least one year). In 2004, the register was adjusted and this resulted in a reduction in the number of foreigners in Germany from 7.3 million to 6.7 million (Opfermann et al. 2006).

Alongside these comprehensive migration statistics, there are also statistical systems for the registration of different groups of immigrants using specific entrance options such as those for ethnic German repatriates, Jewish migrants, family reunification or asylum seekers, short-term labor migrants, highly qualified labor migrants, or students (BAMF 2010).

## 2.2 Databases for research on foreigners and migrants (population stock)

There are three kinds of official databases. Data on population stocks are collected by the local authorities, who adjust the census data of 1987 by adding or subtracting national and international migration flows and natural population events (births and deaths) to or from the population projections of the Federal Statistical Office (*Bevölkerungsfortschreibung*). The population projections contain basic data on the demographic development of the German and foreign population, including statistics on population differentiated according to German and foreign citizenship. We know that the figure of 7.3 million foreigners listed in the population projections is higher than the actual figure. Another shortcoming concerns the very important concept of “foreigners,” according to which German citizens such as the large groups of naturalized migrants and ethnic German repatriates are not identified as migrants. Moreover, new legislation on citizenship includes elements of *Jus soli* for children of foreigners, which will result in a significant decline in the number of “foreigners” over the next few years. This means that the population projections are less and less suitable as a basis for depicting the migrant population.

The BAMF is responsible for the AZR. Data on the foreign population is collected by the local authorities responsible for alien registration and includes the personal details used for administrative purposes such as name, gender, date of birth, date of immigration, country and place of birth, and citizenship. The register has several advantages, the most important of which is the “legal status” parameter, which is useful for differentiating temporary or permanent residents. In addition, unlike the population register, this register permits the identification of first and second generations as well as the duration of stay. The quality of the data was enhanced in 2004 by the clearing up of the AZR, and this resulted in a decrease in the number of foreigners calculated – from 7.3 million to 6.7 million (Opfermann et al. 2006).

A third database for migrant information is the German Microcensus, a mandatory survey of a one percent sample of the German population conducted each year. The last census in Germany was conducted in 1987, and the next one will be in 2011. Until then, the sole official source of information on migrant population households is the Microcensus (Statistisches Bundesamt 2008). The Microcensus is also part of the European Labour Force Survey. Before 2005, migrants were identified by foreign citizenship. To ensure that the figures also covered German migrants, the new concept of “migratory background” was incorporated in the 2005 Microcensus. This concept covers all foreigners, naturalized foreigners, ethnic German repatriates, and immigrants, as well as

their descendants (Statistisches Bundesamt 2008). Under the new definitions, around 15 million persons have a migration background.

Information on religious minorities is not contained in official datasets. The Muslim population in Germany is estimated at 4 million, based on a survey of persons with migration background (Haug et al. 2009).

### 2.3 Databases for integration research

There is an even broader range of options in the field of migrant integration. All the datasets including those providing information on nationality or migration background can be used for social and economic research on integration issues.

At the national, regional, state, and municipal levels, various models are in place for the classification of aggregate data collected for administrative purposes and used in the implementation of an integration monitoring system. Various datasets exist on structural integration, such as education statistics, employment statistics from the Federal Employment Agency (BA, *Bundesagentur für Arbeit*), the datasets of the Institute for Employment Research (IAB, *Institut für Arbeitsmarkt- und Berufsforschung*), and statistics on social security recipients and crime statistics, which also cover German repatriates in some of the states or *Länder*.

The Microcensus is the largest official microdata file and one of the most important resources for integration research. It primarily supports research on the structural integration of migrants and, since 2005, the analysis of persons with a migration background. Researchers interested in aspects other than structural integration are likely to conduct their own surveys or to turn to unofficial data sources.

A second important dataset for the analysis of different aspects of integration is the German Socio-Economic Panel (SOEP, *Sozio-oekonomisches Panel*), a representative longitudinal study of private households conducted annually since 1984 by the German Institute for Economic Research (DIW Berlin, *Deutsches Institut für Wirtschaftsforschung*) (Wagner et al. 2007). Migrants are generally included in the former labor migrant household sample, supplemented by samples of ethnic German repatriates and migrants from a variety of countries.

Many institutes and universities are active in the field of integration research. An overview of ongoing projects and publications can be found in the GESIS-IZ research database (2008). Some institutes have also conducted surveys on the largest migrant groups such as the integration survey of the Federal Institute for Population Research (BiB, *Bundesinstitut für Bevölkerungsforschung*) (Haug and Diehl 2005), the Foreigner Survey of the German Youth Institute (DJI, *Deutsches Jugendinstitut*) (Weidacher 2000) or the representative survey “Selected Groups of Migrants in Germany” (RAM)



(Babka von Gostomski 2008). One current trend is to include a sample of Turkish migrants parallel to an ongoing study, such as in the study on participation and volunteering (Halm and Sauer 2007). Another method of research into integration is to survey migrants in studies on pupils conducted in schools, like the PISA study (Stanat and Christensen 2006), the studies of the Institute for Interdisciplinary Research on Conflict and Violence (IKG, *Institut für interdisziplinäre Konflikt- und Gewaltforschung*) or the Criminological Research Institute of Lower Saxony (KfN, *Kriminologisches Forschungsinstitut Niedersachsen*).

### **3. Future developments in Germany**

#### *3.1 Data collection and data provision*

The most important actors in the field of official statistics are the Federal Statistical Office, the Statistical Offices of the *Länder*, and the BAMF. As in other demographic fields, the most important future developments for migration research are, first, the 2011 Census and, second, the plans currently under discussion to compile a central population register. Both developments will result in a revision of the official number of foreigners.

Private research institutes play an important role in data collection and are involved in almost every large empirical research project (Mohler and Rosenblatt 2008). There is a trend toward including a migrant sample in new survey projects, as in the National Educational Panel Study (NEPS) at the University of Bamberg, or in the Panel Analysis of Intimate Relationships and Family Dynamics (pairfam) at the University of Bremen, or in the Generations and Gender Survey (GGS) of the BiB. The most appropriate methodology for research on the migration process draws from combined surveys in both the country of origin and destination, for example on Mexican-US migration and the Polish migration to Germany (Massey et al. 2008), or the SOEP study on emigrants (Schupp et al. 2008).

In view of the growing migrant population with German citizenship, another trend that can be seen is the use of a name-based sampling method (onomastics, see Humpert and Schneiderheinze 2000; Haug and Diehl 2005; Haug et al. 2009). Another sampling method is based on the birthplace of migrants (topomastics), see for example studies on ethnic German repatriates in the local population register (Haug and Sauer 2007; Salentin 2007).

### 3.2 Data usage and data access

For data protection reasons, there is no free access to local population registers and to the AZR. The registry data is used for administrative purposes, and options for statistical analysis are strictly regulated by law. Aggregate data is available at each office responsible for the register. The local statistical offices give access to population register data for research purposes, which is especially the case for registration-based sampling in survey research.

Aggregate migration and population data can be found in the annual publications of the Federal Statistical Office. Other migration data is published by the BAMF or the BA. Access to the data files can be requested at the local statistical offices or the Statistical Offices of the *Länder*. The Federal Statistical Office publishes data on migration at local level (*Statistik lokal*). Population datasets on the district level are available at the Federal Statistical Office (*Statistik regional*) and through the Federal Office for Building and Regional Planning (BBR, *Bundesamt für Bauwesen und Raumordnung*; INKAR PRO offers up-to-date regional monitoring including future projections). There are many options for researchers who want to access aggregate data on structural integration. Statistics are published by the relevant authorities and by the Federal Statistical Office. For scientific purposes, researchers can access the 2005 Microcensus file at the Research Data Centers of the Federal Statistical Office and the Statistical Offices of the German *Länder*. Access is also possible through the German Microdata Lab at the Leibniz-Institute for the Social Sciences (GESIS, *Leibniz-Institut für Sozialwissenschaften*) in Mannheim. Access to the Microcensus is exemplary, and this option is widely used by researchers.

The most important actors in the research field are the following: the European Forum on Migration Studies (efms), BiB, DIW, DJI, the Migration Research Group of the Hamburg Institute of International Economics (HWWI, *Hamburgisches WeltWirtschaftsinstitut*), IAB, IKG, the Institute for Migration Research and Intercultural Studies (IMIS, *Institut für Migrationsforschung und Interkulturelle Studien*), the Institute for the Study of Labor (IZA, *Forschungsinstitut zur Zukunft der Arbeit*), the Mannheim Center of European Social Research (MZES, *Mannheimer Zentrum für europäische Sozialforschung*), the Social Science Research Center in Berlin (WZB, *Wissenschaftszentrum Berlin für Sozialforschung*), and the BAMF Research Group. A new research institution is the Advisory Board of German Foundations on Integration and Migration (Sachverständigenrat deutscher Stiftungen für Integration und Migration). Moreover, several universities conduct empirical projects on migration or integration.

Research institutes collecting datasets have their own release strategy. Data surveyed for policy-related reasons is generally not available to re-

searchers. The exceptions are the SOEP data, which can be directly ordered from the DIW, or numerous files of the Research Data Center of the BA at the IAB, which are prepared for scientific purposes. The GESIS Data Archive lists other migrant surveys which can be ordered. The most important of these is the repeated MARPLAN survey on foreigners in Germany (latest dataset: 2002). Since 1991, the German General Social Survey (ALLBUS, *Allgemeine Bevölkerungsumfrage der Sozialwissenschaften*) conducted every second year by GESIS has also included foreigners if they speak German. Even if the migrant subgroup is too small to use in integration research, the ALLBUS is an important dataset for the analysis of attitudes among migrant populations.

#### **4. Future developments: European and international challenges**

International migration researchers are faced with differences in key concepts, data collection methods, and databases. As critics have long pointed out (Lederer 2004; Sachverständigenrat 2004), the German system of statistics does not conform to UN recommendations for conceptualizing international migration statistics (determination of long-term and short-term migrants) or international statistics for population stocks (foreign-born persons) (UN 1998; 2007). The established common rules for the collection of statistics by the Member States of the European Community refers to the concept of long-term resident, which is not applied in Germany (EU 2007). The aim of several follow-up international projects (COMPSTAT, THESIM, PROMINSTAT) has been to collect metadata for the comparison of migration and integration of migrants in the EU (Poulain et al. 2006; Fassmann et al. 2008). Clearly, attempts to harmonize migration statistics in the EU have not been very successful to date, making comparisons difficult. The variations reflect differences in the definition of what constitutes a migrant and differences in the data sources – surveys or administrative records (Thierry 2008). However, the duration of stay, the country of birth or the migration background are added to nationality of migrants in some German official statistics (BAMF 2010).

In the field of internationally comparative research on integration, the SOEP is one of the most suitable resources. The SEOP contributes the German section of the household panel of the EU (ECHP). Other datasets that contain information on migrants are the European Statistics on Income and Living Conditions (EU-SILC), the EU Labour Force Survey, the OECD/SOPEMI dataset, the Eurobarometer, and the European Social Survey (see Kahanec and Zimmermann 2008). International projects focusing on

migrants, such as EFFNATIS (efms) or PIONEUR (GESIS, formerly ZUMA), incorporate data on migrants in a comparative perspective. The research program of the IMISCOE network (International Migration, Integration and Social Cohesion), comprised of 23 European research institutes, includes a number of comparative projects, for example the TIES project on the Integration of the European Second Generation.

## 5. Conclusions and recommendations

Generally speaking, the research infrastructure for migration and integration in Germany is very comprehensive. Nevertheless, there are several problems that need to be addressed.

Relating to official migration and integration statistics:

- (1) When using data on migration and migrant population, researchers must deal with the under representation of certain groups of international or internal migrants due to non-registered migration. Studies on irregular migrants can provide an important supplement to knowledge about migration (Cyrus 2008). A more serious problem is the substantial difference between the number of foreigners listed in the population update and the Microcensus on the one hand, and the number registered in the AZR on the other. The lawful adjustment of the local resident registration offices and the central register of foreigners will improve the data quality in the future. The problem may also be tackled by the *2011 Census* and the population update revision. The census will be a new basis for the extrapolation of the migrant population in the Microcensus and also for the weighting of migrants in surveys. Researchers and private research institutes should take heed of these results for present and future projects on migrants.
- (2) The creation of a *central population register* additionally to the local population registers of the municipalities would be another way of improving migration statistics. Researchers should have ensured access to such a central register for the analysis and sampling of foreigners and foreign-born persons – which is currently the case with local population registers. A central population register is cost- and time-effective, and would be a great improvement for survey research. Due to the decentralized register, a nationwide population sampling procedure, for example the ALLBUS, takes five or six months (Babka von Gostomski and Pupeter 2008).

- (3) The *concept of migration background* is now widely accepted. Partial replacement of the term “foreigner” by the concept of “migration background” in official statistics represents a much-needed improvement, that more accurately reflects the reality of the population. In keeping with the recommendations of the working group on Coordinated Household and Population Statistics (HHSTAT) of the Union of Communal Statistical Information Systems (KOSIS-Verbund), statisticians analyze local population registers for persons with migration backgrounds (Härle 2004; Bömermann et al. 2008). There are also approaches that have been developed to make use of the concept of migration background in education and labor-market statistics. Yet there is no agreement or common use of the concept in several key respects, such as the migration status of third-generation migrants, or children who have only one parent with a migration background. It is still not possible to definitively identify ethnic German repatriates in the Microcensus 2007 (Seifert 2008). Furthermore, it is clear in the literature that researchers use slightly different definitions of migration background. Variations in how this concept is operationalized complicate the interpretation of results, and researchers should carefully consider the implications of their specific definitions of migration background.
- (4) In general, data for *international comparative research* on migration is unsatisfactory. Unlike other countries, the concept of ethnicity is not used in official German statistics (see Schönwälder et al. 2008 for an overview of ethnicity measurements in empirical research). The extension of the “foreigner” and “foreign-born” concepts and the inclusion of “descendants of foreign-born” is suggested for the 2011 Census in the EU (UN 2006). Germany will participate in the EU-wide population census scheduled for 2011. After a long-lasting debate on the use of the concept of migration background, recommendations of experts and advice of researchers have been taken. The questionnaire of the 2011 Census will entail questions on migration background. The Census data will be a solid foundation for extrapolation of persons with migration background in the Microcensus in the future and facilitate comparative research on international migration as well as studies on integration of migrants.
- (5) German migration statistics do not include information on the duration of stay, so the concept of *short-term/long-term migration* is not practicable. The AZR supports an analysis of long-term foreign migrants but excludes German citizens. A new law on population statistics was passed in 2008, stipulating that immigrants must be asked about their date of emigration. In this way, long-term emigrants who stay out of the country for longer than one year can be identified when they return to Germany,

and circular migration can be identified. This is an improvement for German migration statistics, but a more comprehensive solution for the analysis of long-term migration is preferable. A central population register would enhance the data situation.

- (6) The implementation of the “*foreign-born*” concept, as it stands now, is inadequate. This problem will be addressed by entering the country of birth of foreigners in the local registers. The new rule is applied for foreigners registering as from 2009 on. This solution represents a real improvement in population statistics, since the current practice is to record only the place of birth.

Relating to empirical migration and integration research:

- (7) In order to improve the validity of social research and social structure analysis, migrants should, in general, be included in *representative population surveys*. One reason for the implicit or explicit exclusion of migrants is that first-generation migrants in particular frequently do not have sufficient German language knowledge to participate in a survey. Translations of questions should thus be available for interviewers, at a minimum in the most prevalent migrant languages, Turkish and Russian.
- (8) The accessible datasets of the Microcensus, the SOEP, or the IAB Employment Sample are adequate for integration research that deals with large migrant groups, and there is still untapped potential for analysis in this area. But even where population surveys like the Microcensus cover the full range of immigrant groups living in Germany, subgroup analyses soon run into case number problems, since data protection regulations make research into smaller migrant groups or small regions difficult. In order to acquire an adequate case number of migrants for analysis, we recommend that *specific large-scale survey studies be supplemented by a migrant sample*. Some possible examples are the Network for integrated European population studies (NIEPS), pairfam, or GGS. The advantage of this is that it enables comparisons between non-migrants and migrants. The majority of projects based on this kind of research design focus on Turkish migrants; however, the number of immigrants from Russia and other CIS countries is even larger.
- (9) Some of the important questions in migration and integration research cannot be adequately studied through the available general surveys. Complex questions on migration biography, for example, refer only to migrants and therefore are not included in the ALLBUS. Also there is a lack of data in the Microcensus on indicators of emotional integration, on attitudes or religion. Instruments for the analysis of particular constraints on and resources available to migrants are important for the

analysis of the causal mechanisms of integration processes. *Adequate instruments* must be developed for these topics in conducting empirical research on migration and integration, and to implement these instruments into large studies.

- (10) Most research results in the German context provide information about the integration of the large and very important group of Turkish migrants. Over and above questions about integration, research has to face the challenges presented by *new groups of migrants*, such as the large and rapidly growing group of Polish migrants or the groups with a unique profile like Jewish or Vietnamese migrants or refugees. One possible way to cope with this problem is to conduct special surveys for these migrant groups. The sampling of small migrant groups in a nationwide context, however, is more complicated. Difficulties arise with the sampling of migrant groups comprised of German citizens – such as ethnic German repatriates and naturalized persons. The identification of the minority population using name-based sampling methods is the preferred option in all these cases. A challenge for future research, therefore, is the development and methodological assessment of new sampling methods based on surnames or place of birth (onomastics and topomastics).
- (11) For the analysis of causal mechanisms in the integration process, more panel studies like the SOEP are needed. Therefore, the implementation of a *longitudinal migrant study* in Germany is in the discussion stage.
- (12) Access to most of the important datasets is possible by subscription (SOEP, ALLBUS) or by the Research Data Centers (Microcensus, IAB datasets). Yet, even if access to official data or to datasets of other research institutes can be granted upon request on a case-by-case basis, general access can represent a serious impediment for researchers.
- (13) Internationally comparable datasets on migration are rare and migration research is in most cases restricted to the country of destination. With the growing importance of *circular migration*, however, the collection of data in the migrant country of origin and/or on the biographies of transnational migrants are becoming wide fields for future research.
- (14) Since data collection is typically delegated to *private-sector agencies*, a competent partner is needed for the complex field work and ensurance of high data quality (Rosenblatt 2008; Mohler and Rosenblatt 2009). Energy and resources need to be invested in the continuous improvement of migrant survey methodology in cooperation with researchers and private research institutes.

## References:

- Babka von Gostomski, C. (2008): Türkische, griechische, italienische und polnische Personen sowie Personen aus den Nachfolgestaaten des ehemaligen Jugoslawien in Deutschland. BAMF Working Paper Nr. 11.
- Babka von Gostomski, C. and Pupeter, M. (2008): Zufallsbefragung von Ausländern auf Basis des Ausländerzentralregisters. Methoden – Daten – Analysen 2/2008, 149-177.
- BAMF (Bundesamt für Migration und Flüchtlinge) (2010): Migrationsbericht 2008. Nuremberg.
- Bömermann, H./Rehkämper, K. and Rockmann, U. (2008): Neue Daten zur Bevölkerung mit Migrationshintergrund in Berlin zum Stand 31.12.2007. Zeitschrift für amtliche Statistik Berlin Brandenburg 3, 20-28.
- Boos-Nünning, U. and Karakasoglu, Y. (2006): Viele Welten leben. Münster.
- Boswell, C. and Mueser, P. (Eds.) (2008): Journal of Ethnic and Migration Studies. Special Issue: Economics in Migration Research: Towards Disciplinary Integration? 34 (4).
- Brettfeld, K. and Wetzels, P. (2007): Muslime in Deutschland. Berlin.
- Cyrus, N. (2008): Country report Germany: Undocumented Migration Counting the Uncountable. Data and Trends across Europe. HWWI. Hamburg.
- Diefenbach, H. and Weiß, A. (2006): Menschen mit Migrationshintergrund. Datenerfassung für die Integrationsberichterstattung. Gutachten für das Statistische Amt & die Stelle für interkulturelle Arbeit der Landeshauptstadt München. Munich.
- Diehl, C. and Haug, S. (2003): Assessing migration and integration in an immigration hesitant country: the sources and problems of data in Germany. Studi Emigrazione/International Journal of Migration Studies 40 (152), 747-771.
- Esser, H. (1980): Aspekte der Wanderungssoziologie. Darmstadt.
- Esser, H. (2008): Assimilation, ethnische Schichtung oder selektive Akkulturation? Neuere Theorien der Eingliederung von Migranten und das Modell der intergenerationalen Integration. Kölner Zeitschrift für Soziologie und Sozialpsychologie, Sonderheft 48, 81-107.
- EU (2007): EU regulation (EC) No 862/2007 of the European Parliament and the council, Community statistics on migration and international protection and repealing Council Regulation (EEC) No 311/76 on the compilation of statistics on foreign workers.
- Fassmann, H./Reeger, U. and Sievers, W. (2008): Statistics and Reality: Concepts and Measurements of Migration in Europe (IMISCOE Reports). Amsterdam.
- Filsinger, D. (2008): Bedingungen erfolgreicher Integration: Integrationsmonitoring und Evaluation; Expertise im Auftrag der Friedrich-Ebert-Stiftung. Bonn.
- GESIS-IZ Sozialwissenschaften und BAMF (2008): Migration und ethnische Minderheiten, SoFid. Bonn.
- Granato, N. (2003): Ethnische Ungleichheit auf dem deutschen Arbeitsmarkt. Opladen.
- Halm, D. and Sauer, M. (2007): Bürgerschaftliches Engagement von Türkinnen und Türken in Deutschland. Wiesbaden.



- Härle, J. (2004): Personen mit Migrationshintergrund. *Stadtforschung und Statistik* 1/2004, 16-18.
- Haug, S. (2002): Familienstand, Schulbildung und Erwerbstätigkeit junger Erwachsener. Erste Ergebnisse des Integrationssurveys des BiB. *Zeitschrift für Bevölkerungswissenschaft* 27 (1), 115-144
- Haug, S. (2005): Die Datenlage im Bereich der Migrations- und Integrationsforschung. BAMF Working Paper Nr. 1.
- Haug, S. and Diehl, C. (2005): Aspekte der Integration. Wiesbaden.
- Haug, S. and Sauer, L. (2007): Zuwanderung und Integration von (Spät-)Aussiedlern. BAMF Forschungsbericht Nr. 3.
- Haug, S./Stichs, A. and Müssig, S. (2009): Muslim Life in Germany. A study conducted on behalf of the German Conference on Islam. Research Report No. 6. Nuremberg.
- Huinink, J. and Kley, S. (2008): Regionaler Kontext und Migrationsentscheidungen im Lebensverlauf. *Kölner Zeitschrift für Soziologie und Sozialpsychologie, Sonderheft 48*, 162-184.
- Humpert, A. and Schneiderheinze, K. (2000): Stichprobenziehung für telefonische Zuwandererumfragen. Einsatzmöglichkeiten der Namensforschung. *ZUMA Nachrichten* 24 (47), 36-64.
- Kahanec, M. and Zimmermann, K. (2008): Migration and Globalization: Challenges and Perspectives for the Research Infrastructure, RatSWD Working Paper No. 51.
- Kalter, F. (Ed.) (2008a): Migration und Integration. *Kölner Zeitschrift für Soziologie und Sozialpsychologie, Sonderheft 48*.
- Kalter, F. (1997): Wohnortwechsel in Deutschland. Opladen.
- Kalter, F. (2003): Stand und Perspektiven der Migrationssoziologie. In: Orth, B. et al. (Eds.): *Soziologische Forschung. Stand und Perspektiven*. Opladen.
- Kalter, F. (2008b): Einleitung. Stand, Herausforderungen und Perspektiven der empirischen Migrationsforschung. *Kölner Zeitschrift für Soziologie und Sozialpsychologie, Sonderheft 48*, 11-36.
- KGSt (Kommunale Gemeinschaftsstelle für Verwaltungsmanagement) (2006): Integrationsmonitoring. Materialien Nr. 2/2006. Cologne.
- Kristen, C. (2005): School choice and ethnic school segregation. Primary school selection in Germany. Münster.
- Lederer, H. (2004): Indikatoren der Migration. Bamberg.
- Massey, D./Kalter, F. and Pren, K. (2008): Structural Economic Change and International Migration from Mexico and Poland. *Kölner Zeitschrift für Soziologie und Sozialpsychologie, Sonderheft 48*, 134-161.
- Mehrländer, U./Ascheberg, C. and Ueltzhöffer, J. (1996): Situation der ausländischen Arbeitnehmer und ihrer Familienangehörigen in der BRD. Repräsentativuntersuchung '95. Berlin.
- Mohler, P. and Rosenblatt, B. von (2009): Infrastructure for large-scale survey measurement: cooperation of academic research and private sector agencies. RatSWD Working Papers No. 69.
- Nauck, B. (2008): Akkulturation: Theoretische Ansätze und Perspektiven in Psychologie und Soziologie. *Kölner Zeitschrift für Soziologie und Sozialpsychologie, Sonderheft 48*, 108-133.

- Opfermann, H./Großbecker, C. and Krack-Roberg, E. (2006): Auswirkungen der Bereinigung des Ausländerzentralregisters auf die amtliche Ausländerstatistik. *Wirtschaft und Statistik* 5, 480-494.
- Poulain, M./Perrin, M. and Singleton, A. (2006): THESIM: Towards Harmonised European Statistics on International Migration. Louvain.
- Pries, L. (1997): Einleitung. Neue Migration im transnationalen Raum. *Transnationale Migration. Soziale Welt, Sonderband 12*, 15-46.
- Rosenblatt, B. von (2008): Die Rolle der Umfrageinstitute in der sozialwissenschaftlichen Dateninfrastruktur. RatSWD Working Paper No. 36.
- Sachverständigenrat für Zuwanderung und Integration (2004): Migration und Integration – Erfahrungen nutzen, Neues wagen. Jahresgutachten 2004. Nuremberg.
- Salentin, K. (2007): Die Aussiedler- Stichprobenziehung. *Methoden – Daten – Analysen* 1 (1), 25-44.
- Salentin, K. and Wilkening, F. (2003): Ausländer, Eingebürgerte und das Problem einer realistischen Zuwanderer-Integrationsbilanz. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 55 (2), 278-298.
- Schönwälder, K./Baykara-Krumme, H. and Schmid, N. (2008): Ethnizität in der Zuwanderungsgesellschaft Deutschland: Zur Beobachtung ethnischer Identifizierungen, Loyalitäten und Gruppenbildungen. SOEB Arbeitspapier 1. Forschungsverbund Berichterstattung zur sozioökonomischen Entwicklung der Bundesrepublik Deutschland.
- Schupp, J./Siegel, N./Erlinghagen, M./Stegmann, T. and Wagner, G.G. (2008): Leben außerhalb Deutschlands: eine Machbarkeitsstudie zur Realisierung von Auslandsbefragungen auf Basis des Sozio-oekonomischen Panels (SOEP). SOEP Papers No. 120.
- Seifert, W. (2008): Aussiedlerinnen und Aussiedler – neue Erfassungsmöglichkeiten und sozioökonomisches Profil. *Statistische Analysen und Studien*, Band 53, 11-23.
- Siegert, M. (2006): Integrationsmonitoring – State of the Art in internationaler Perspektive. Nuremberg.
- Stanat, P. and Christensen, G. (2006): Schulerfolg von Jugendlichen mit Migrationshintergrund im internationalen Vergleich. Eine Analyse von Voraussetzungen und Erträgen schulischen Lernens im Rahmen von PISA 2003. Berlin.
- Statistisches Bundesamt (with WZB and ZUMA) (2006): Datenreport 2006. Bonn.
- Statistisches Bundesamt (2008): Bevölkerung mit Migrationshintergrund. Ergebnisse des Mikrozensus 2007. Fachserie 1. Reihe 2.2. Wiesbaden.
- Thierry, X. (2008): Towards a harmonization of European Statistics on International Migration. *Population & Societies* No. 442.
- UN (1998): Recommendations on Statistics of International Migration. Rev. 1. New York.
- UN (2006): Conference of European Statisticians. Recommendations for the 2010 Censuses of Population and Housing. New York.
- UN (2007): Principles and Recommendations for Population and Housing Censuses, Rev. 2. New York.
- Venema, M. and Grimm, C. (2002): Situation der ausländischen Arbeitnehmer und ihrer Familienangehörigen in der BRD. Repräsentativuntersuchung 2001. Offenbach.

- Wagner, G.G./Frick, J.R. and Schupp, J. (2007): The German Socio-Economic Panel Study (SOEP) – Scope, Evolution and Enhancements. *Schmollers Jahrbuch* 127 (1), 139-169.
- Wagner, M. (1989): *Räumliche Mobilität im Lebensverlauf. Eine empirische Untersuchung zu sozialen Bedingungen der Migration.* Stuttgart.
- Weidacher, A. (2000): *In Deutschland zu Hause.* Opladen.
- Windzio, M. (2004): Kann der regionale Kontext zur „Arbeitslosenfalle“ werden? Der Einfluss der Arbeitslosigkeit auf die Mobilität zwischen regionalen Arbeitsmärkten in Westdeutschland. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 56 (2), 257-278.
- Worbs, S. and Friedrich, L. (2008): Integrationsberichterstattung in Deutschland. Eine Bestandsaufnahme. *Sozialwissenschaften und Berufspraxis* 2/2008, 250-269.



### 1.3 Internal Migration

*Andreas Farwick*

Contact:

Andreas Farwick  
Ruhr University Bochum  
Geography Department  
Universitätsstraße 150  
44780 Bochum  
Germany  
e-mail: andreas.farwick[at]ruhr-uni-bochum.de

## **Abstract**

Research on internal migration covers a wide range of issues that pertain to the reasons for moving, the distance and direction of movement within a country, and the process of decision making involved in undertaking these moves. Given the rich field of relevant research objectives and the substantial developments in migration theory, it is clear that the availability of a broad set of data that includes detailed information on various aspects of life is one of the key factors in ensuring continued progress in the analysis of internal migration and its development. The available official aggregated data are useful for descriptive structural analyses; however, they are limited in their ability to explain causal relations. The same holds true for cross-sectional data. Some of the longitudinal datasets discussed consist of retrospectively collected event history data that are not suitable for acquiring essential information about the attitudes and psychological states of the respondents over time. Several prospective longitudinal survey data do not represent essential aspects of internal migration. Data should at least include information on the place of residence (on the smallest possible spatial level), typologies pertaining to the characteristics of the place of residence, any change in residence, reasons for a move, the plan to migrate, the type of dwelling and the neighborhood, as well as on commuting.

Keywords: internal migration, regional migration, migration theory, official data, cross-sectional data, longitudinal data

## **1. Research objectives**

The main fields of research in the study of internal migration can be addressed by posing the following basic questions: (a) who moves, (b) why do they move, (c) from what origin, (d) to what destination, (e) how are the decision-making processes involved in the move determined, and finally (f) how does this process change over time? These essential issues structure the field of research.

Issues a and b refer to the reasons and motives to change residence. One can make some rough distinctions between education-related movers, workplace-related movers, housing-related movers, and retirement-related movers (Gatzweiler 1975). This classification scheme runs parallel to specific stages in the life course (Rossi 1955) and can be related to different age groups: education-related movers (age 16 to 20 years), workplace-related movers (age 21 to 34 years), housing-related movers (age 25 to 49 years) and retirement-related movers (age 49 and above). Whether this classification of movers as an explanatory characteristic will remain adequate over time is an open question.

With regard to the origin and destination of movers (questions c and d), distinctions can be made between short- and long-distance migration and moves between different types of regions (e.g., rural to urban and vice versa). These migration patterns are again to some extent related to the reasons for the move. Education-related movers mostly stem from peripheral rural regions with an unattractive and relatively undifferentiated range of educational facilities. Workplace-related movers stem from peripheral or declining old industrial regions with shrinking opportunities for qualified workers (“rust belts”) to the metropolitan centers of growth industries (“sun belts”). In Germany this workplace-related interregional migration has taken on a major significance since the early 1970s in the form of a North to South shift (Friedrichs et al. 1986; Windzio 2004) and – following the reunification of Germany – in the form of a massive East to West movement, especially in the first two years (Büchel and Schwarze 1994; Burda 1993; Wagner 1992; Windzio 2007; 2009).

Housing-related migration patterns are predominantly intra-regional or intra-urban. One of the major intra-regional migration patterns is the process of suburbanization, which began in the early 1960s. Particularly during the 1970s, increasing family income, improved transportation systems, and public incentive programs that encouraged individual housing, led to a first wave of the population shift away from the central cities. A second wave of suburbanization took place at the end of the 1980s. Increasing population densities and the extension of suburban areas pushed the new suburbanites further and further into the urban peripheries, causing substantial urban sprawl (Bleck and Wagner 2006). Due to the recent rebirth of inner-city housing for broad sections of the population, there is now a notable trend towards reurbanization (Brühl et al. 2005).

Currently, retirement-related movement in Germany does not have the magnitude that it has in the United States or in France, but it is becoming more and more significant. Retirement seekers notably favor regions with attractive landscapes, such as the northern foothills of the Alps (*Alpenvorland*) (Friedrich 1995).

Over time these population shifts will cumulatively increase regional disparities and may have substantial negative consequences for demographic and economic development, since all of the types of migration described are highly selective with respect to age, gender, and economic status. The migration of younger individuals, for example, results in the massive aging of peripheral regions, with consequences for the natural reproduction of the population. Population losses in East Germany and in the traditionally industrial areas of West Germany, combined with an ongoing suburbanization cause a considerable shrinking of central cities (Eichstädt-Bohlig et al. 2006). Furthermore, extensive out-migration of highly skilled labor from decreasing to prosperous regions (“brain drain”) leads to a decrease in the



human capital necessary for further development (Friedrich and Schultz 2008). Competing for highly skilled employees, some city business development agencies profile themselves as an attractive destination for qualified workers by stimulating – according to the thesis of Florida (2004) – a tolerant and cultural diverse climate in their regions.

The selectivity of migration patterns also raises a problem with regard to intra-urban migration. While well-educated, high-income city dwellers (“yuppies” and “dinkies”) tend to rent or buy apartments or houses in the renovated and upgraded nineteenth-century inner city residential areas (thus contributing to a process of gentrification), an increasing number of lower income groups must move to the more run-down sections of the traditional working-class areas, or to peripheral public housing estates at the outskirts of the central cities to find less expensive rentals. This pattern of selective intra-urban migration causes a high degree of residential segregation and leads to the rise of poverty areas in which the social problems of the residents – due to negative neighborhood effects – accumulate (Farwick 2001).

While the decision to move can be partially explained by a typology of reasons and differences in opportunity structures (supply with infrastructure, labor market, housing market, climate, landscape, etc.), it remains difficult to explain why some people move and others do not. This issue is linked to the questions introduced at the opening of this discussion, that pertain to decision-making processes. On the one side, objective individual characteristics (age, gender, educational attainment, occupational or family related conditions, as well as housing conditions) are important factors in explaining these processes. On the other side, subjective factors like motives, information, and the evaluation of the situation, play an important role too. According to Kalter (1997), the decision to move can be divided into three stages: the idea of moving, the plan to move, and the actual move. The challenge is thus to explain the factors that determine each stage in the overall decision-making process. The complexity of analyzing the decision-making process becomes apparent when one considers that it is embedded in the life course and therefore related to many other events that take place during the lifetime (Wagner 1989).

Moreover – like every social action – the decision to change residence is framed by the social, political, and economic conditions of society. Since these social conditions change continuously over time, research questions seek to explain how different migration processes refer to these ongoing changes in the social environment.

Theoretical concepts approach the investigation of the phenomenon of internal migration on both the macro- and on the micro-level. Based on Ravenstein’s classic “Laws of Migration” (1885/1889, reprint 1995), which emphasize the significance of the distance between the origin and destination in migration as a means of estimating population flows – the gravity model is

the most important concept explaining internal migration patterns at the macro-level (Birg et al. 1993). Introducing other regional characteristics in addition to population size and distance, however, would extend this model. For example, neoclassical economic theories stress the role of regional income and job vacancy differentials particularly for an explanation of interregional migration patterns (Todaro 1969). With regard to intra-urban migration, concepts focusing on structures of supply and demand on the regional housing market are of particular relevance (Farwick 2001).

A major shortcoming of migration theories on the macro-level is that they cannot explain exactly how the decision to change residence is affected by regional characteristics. In this context, Lee (1972) outlined the impact of intervening obstacles. He argued that diverse variables – such as distance, physical and political barriers, and having dependents – could impede or even prevent migration.

Sjaastads (1962) seminal work considers migration as a particularly important investment decision in human capital. In the simplest model of wealth maximization, the fixed costs of moving are balanced against the net present value of earnings streams available in the alternative location. Furthermore, the social psychological approach adopted by Wolpert (1965) characterizes migration as a form of individual or group adaptation to perceived changes in environment.

A synthesis of different approaches to explain migration behaviors has given rise to the “value-expectancy model” (De Jong and Fawcett 1981). According to this, the decision to move is based on a specification of the personally valued goals that might be achieved by moving (or staying) and the perceived linkage, in terms of expectancy, between migration behavior and the attainment of these goals in alternative locations. Kalter (1997) enhanced this model in three ways: by incorporating the cost-benefit calculus of households, by accounting for the tendency toward inertia, and by integrating the problems of constraints and facilitators of the environment. Consequently, the decision to move has to be operationalized as a significant part of the life course characterized by a high degree of interdependence with other areas of life (Huinink and Kley 2008; Wagner 1989).

The described research objectives and theoretical developments show that studies on internal migration remain on the scientific frontier. Particularly, attention must be given to the development of theoretical models and empirical methods able to connect the decision-making process involved in migration to the complexity of events in a life-course perspective. We need more insight into the process of considering a change in residence or, alternatively, into the choice to commute – even over long distances – increasingly in the form of multi-local living arrangements.

Investigating migration as a combined and complex decision-making process influenced by a variety of family members is another important

research area. The influence of broader social networks on the decision to move and on the destination of a move must also be considered.

An ongoing methodological challenge for studying migration decisions is the problem of self-selection, which has been noted by Borjas (1987). Since characteristics that influence wages also influence migration, specific methods – such as those described in Heckman (1979) for example – are needed to deal with this bias (Massey and Espinosa 1997; Windzio 2007; see also the report by Kahanec and Zimmermann on migration and globalization in this publication).

## **2. Status quo: Databases and accessibility**

Over the last few decades we have seen considerable theoretical and methodological progress in this area of research. Yet to render these developments fruitful and to adequately meet the relevant research objectives, a rich pool of data is needed that can be applied to all kinds of analysis on the macro- or micro-level, with cross-sectional or panel design respectively.

### *2.1 Official statistics*

Data from official statistics are used to describe structures of internal migration and to analyze processes on the macro-level (e.g., Schlömer and Bucher 2001). Data on population flows that result from residential moves are based on registration and deregistration within specific municipalities and are available from the Federal Statistical Office in the form of migration matrixes on different administrative levels, ranging from the federal states (*Bundesländer*) to rural and city districts (*ländliche Kreise* and *kreisfreie Städte*). The Statistical Offices of the German *Länder* provide migration matrixes at the spatial level of cities and communities. In case of many cities, migration matrixes are also available for intra-urban moves.

On an aggregated level these official data differentiate between the individual characteristics of age, gender, nationality, and employment status. The data serve to calculate various descriptive measures of migration, to identify interdependencies between regions, and to adopt gravity models (Birg et al. 1993). Since in the gravity models, the distance between the sources and destinations of movements is used, it would be a substantial improvement if migration matrixes from the Federal Statistical Office could include data on distances between the corresponding regions.

Spatial context information at different spatial levels down to the rural or city districts are available from the Federal Statistical Office and the Statisti-

cal Offices of the German *Länder*. Together these offices provide a data collection called “Regio-Stat-Katalog” which contains a variety of different regional characteristics (Arbeitsgruppe Regionale Standards 2005). The same information is also available on CD-ROM under the label “CD-ROM Statistik Regional.” Data on an even smaller level of the more than 12,000 German cities and communities are provided by a collection called “DVD Statistik Lokal” which is annually updated. Another excellent source of regional data with a broad range of spatial characteristics with respect to different areas of life is a collection published on CD-ROM by the Federal Office for Building and Regional Planning (BBR, *Bundesamt für Bauwesen und Raumordnung*) called “INKAR.”

A source of official data at the individual level is the German Microcensus (Wirth et al. 2005). For research purposes the data can be obtained from the Federal Statistical Office in the form of a Scientific Use File (SUF) that describes the place of residence of the respondents at the level of federal states, or *Länder*, and in form of a typology of communities by population size (*Gemeindegrößenklasse*). The data also contain the ID of the sample district (*Auswahlbezirk*) from which each individual in the sample is drawn. Of importance for internal migration research is information on residential change (since the previous year) and housing conditions. Comprehensive data on commuting to work is available for the years 1996, 2000, and 2004. The data include no information on reasons for a move or the plan to migrate. The Federal Statistical Office is planning to release a Microcensus Regional File that will include regional information at the level of 349 Microcensus districts (MZKR, *Mikrozensus-Kreisregionen*). Unfortunately, this file will not include information about residential changes undertaken since the previous year.

The Microcensus is a rotating panel sample in which every household within the sample district is included for a four-year time period (Lüttinger and Riede 1997). However, because households that change their residence during that period drop out of the sample (Rendtel 2005; see also the report on Family included in this publication), the panel is more or less useless for internal migration research.

Labor migration can be studied using the IAB Regional Sample, a regional version of the IAB Employment Sample provided by the Institute of Employment Research (IAB, *Institut für Arbeitsmarkt- und Berufsforschung*) (e.g., Windzio 2004; 2007; 2009). The data consists of a two-percent subsample of all employees in Germany subject to social insurance contribution, supplemented by information on benefit recipients. The file includes information from 1975 (in West Germany) to 2004 (Drews 2008). The sample covers a continuous flow of data on employment subject to social security as well as on the receipt of unemployment benefits, unemployment assistance, and maintenance allowance. Data include the district number

(*Kreiskennziffer*) of the workplace. They do not provide information on the place of residence. Because of this fact it is not possible to distinguish whether a change in workplace is connected with a residential move or a change in commuting to work. Therefore, the inclusion of the place of residence into the dataset should be taken into consideration. The data could take the form of a Scientific Use File delivered via the Leibniz Institute for the Social Sciences (GESIS, *Leibniz-Institut für Sozialwissenschaften*).

## 2.2 Survey data

One of the most important data for research on internal migration is the German Socio-Economic Panel (SOEP, *Sozio-oekonomisches Panel*) (Wagner et al. 2007), a representative longitudinal study of private households conducted annually by the German Institute for Economic Research (DIW Berlin, *Deutsches Institut für Wirtschaftsforschung*) (Burda 1993; Büchel and Schwarze 1994; Hunt 2004; Jürges 1998; Kalter 1994; Wagner 1992). Regional information about the place of residence is available on different spatial levels down to the German zip code areas (Spieß 2005). Regional typologies (community type, community size) are also available. Since 2004, the information on the place of residence is matched with geographical microdata from MICROM Consumer Marketing. These data – in the form of various MOSAIC typologies – contain information for housing blocks concerning demographic characteristics, housing type, car use, mobility, consumer behavior, social milieus, and purchasing power (Goebel et al. 2007). The SOEP dataset itself includes key indicators like date of move, reasons for move, and the plan to migrate. In addition, the data give information about housing status, quality of dwelling, and neighborhood characteristics. Since the SOEP allows for combining information on all household members, it is possible to apply multi-actor analytical designs. The usefulness of the SOEP dataset is greatly enhanced by this wide variety of structural characteristics as well as the inclusion of attitude indicators.

Another longitudinal dataset is the German Life History Study (GLHS) originally conducted by the Max Planck Institute for Human Development in Berlin (MPIB) and now continued at Yale University. The GLHS comprises the life histories of some 8,500 men and women from twenty selected birth cohorts in West Germany and of more than 2,900 men and women from thirteen selected birth cohorts in East Germany. West Germans born in 1964 and 1971 were interviewed in 1998–99 with a sample size of 2,909 respondents. A follow-up with the 1971 cohort was completed in 2005. The GLHS has an explicit focus on residential and migration history (Wagner 1989; Rusconi 2006). Detailed retrospective life-course information is available for all moves, including reasons to move, housing conditions, type of residential place, and type of neighborhood. Information on intentions to move in the

future is lacking. In the Public Use Files available at GESIS, direct references to places and all open-ended responses have been removed.

The German Youth Institute (DJI, *Deutsches Jugendinstitut*) has conducted the Family Survey that is to some extent useful for migration research. It is a recurring survey of approximately 10,000 respondents that was conducted in an interval of six years (1988, 1994, and 2000). For a subsample of about 2,000 respondents it includes a three-wave panel. Regional information on the place of residence is available on different spatial levels down to the rural or city districts (*Kreise*). Moreover regional typologies of the places of residence are available in the form of the BIK typology (Hoffmeyer-Zlotnik 2005) and – for the third wave in the year 2000 – in form of the MOSAIC typologies from MICROM Consumer Marketing (see SOEP). In addition, the data include information on housing status and characteristics, quality of dwelling, and neighborhood characteristics. The cross-sectional dataset of the year 2000 contains also questions about reasons for leaving, respectively returning to the parental home and reasons for the first three changes of residence since age of 16. The data are available directly via the German Youth Institute website.

A more recent longitudinal dataset is the German Generations and Gender Survey (GGS), an international comparative panel study coordinated by the United Nations Economic Commission for Europe (UNECE) in Geneva. The Federal Institute for Population Research conducts the German part of the survey (Ruckdeschel et al. 2006). The first wave of the GGS was collected in 2005. In 2006 another sample of Turkish migrants was undertaken. Data collection for the second wave began in 2008. The data contain housing characteristics and questions about the intention to change residence. The data can be requested at the Federal Institute for Population Research.

In analyzing mobility patterns of the elderly, the Survey of Health, Ageing and Retirement in Europe (SHARE) (Börsch-Supan et al. 2003) is useful. To date, two waves with respondents aged 50 and over have been conducted in 2004 and in 2006. A third wave is currently in progress. The data include regional information about the place of residence on different spatial levels down to the rural or city districts (*Kreise*). Unfortunately for the German sample, data on residential location are only provided on the level of the *Länder*. Information on the housing situation, a change of the place of residence, and the main reasons for a move are available but information on the plan to migrate does not exist.

The German General Social Survey (ALLBUS, *Allgemeine Bevölkerungsumfrage der Sozialwissenschaften*), a cross-sectional database also provided by GESIS, is not applicable to research in the field of internal migration. Questions about the duration of habitation – in both a specific apartment or house and the location of residence – as well as the distance from the former place of residence, have only been included since 2000. It

would be worthwhile to consider adding additional questions about possible intentions to change residence and the assessment of living conditions at the current domicile.

### 3. Future developments

One very welcome development in gaining access to data for migration studies is the increasing availability of official data via the Internet provided by various public institutions. The Federal Statistical Office together with the Statistical Offices of the German *Länder* offer migration statistics on the level of the *Länder* and rural or city districts respectively through their Internet platform, the “German Regional Database” (“Regionaldatenbank Deutschland”). Moreover, the Lower Saxony Institute for Statistics and Communication Technology (LSKN, *Landesbetrieb für Statistik und Kommunikationstechnologie Niedersachsen*) has made particularly successful efforts to provide comprehensive regional migration data for the state of Lower Saxony that includes data all the way down to the level of the city and community, which are accessible via its system “LSKN Online.” Intra-urban migration data are, for example, provided by the Statistical Office of the city of Bremen (*Statistisches Landesamt Bremen*) through its excellent information system “Bremen on a Small Scale” (“Bremen Kleinräumig”). These examples should encourage other federal states, cities, and communities to offer regional data in a comprehensive way through the Internet.

Another advance in the access of data relating to the study of internal migration has accompanied the establishment of the Research Data Center of the Federal Statistical Office, which provides on-site use of official survey data (e.g., the German Census or Microcensus) and off-site use of different Public or Scientific Use Files. The same holds for the Research Data Center of the Federal Employment Agency at the Institute of Employment Research (IAB, *Institut für Arbeitsmarkt- und Berufsforschung* within the BA, *Bundesagentur für Arbeit*). These efforts need not only to be continued but also expanded.

Since the possibilities for in-depth analysis on the causal relations of migration through the use of official data are currently very limited, survey data will continue to play a major role. One future challenge in the field of internal migration research is to further the understanding of interdependencies between migration decisions and regional opportunity structures in the context of the life course. With regard to this question, Huinink and Kley (2008) stress that the significance of contextual impacts is strongly related to the aims and demands of actors in specific stages of their life course, a fact that has only been given rudimentary theoretical and empirical

analysis. Studies that want to address these issues require comprehensive longitudinal datasets including information on the place of residence that can be combined with an adequate variety of regional characteristics. Positive developments in this direction can be seen by the efforts of the DIW Berlin to make small-scale regional information of the SOEP available for analyses and to link them with spatial information from other datasets (see above).

#### **4. Conclusion and recommendations**

Given the range of research objectives and developments in migration theory, it is apparent that the availability of a broad set of data, including detailed information on various aspects of life (in particular educational and occupational biographies as well as changes in household structure), combined with information on the regional structure of the place of residence, is one of the key factors for ongoing progress in research on internal migration. The datasets described here are more or less sufficient to meet these demands.

Official aggregated data are particularly useful for descriptive structural analyses. As far as possible they should be made accessible via the Internet. For explaining causal relations, however, the value of aggregated data is limited. Therefore, cross-sectional survey data and especially longitudinal datasets are needed.

Among the cross-sectional survey data described, the German Microcensus is of immense importance – not least because of its huge sample size. Its value could be improved by collecting information on the reasons behind a given move and intentions to move. It is therefore strongly recommended that the Microcensus Regional File include information on residential change (since the previous year), reasons for a move, and the plan to migrate. In terms of migration studies, the usefulness of the IAB Regional Sample, the regional version of the IAB Employment Sample, would be greatly enhanced by including information on the place of residence.

Some of the described longitudinal datasets (the GHLS or the DJI Family Survey) consist of retrospectively collected event history data. The problem with this data lies in the inability of the surveys to collect information about the attitudes and psychological states of the respondents over time. Therefore, they do not provide characteristics like the subjective evaluation of what opportunities exist in the residential environment, or the emotional closeness to the place of residence, which are both highly relevant factors in migration intentions and actual migration. In light of these problems, the continuation and optimization in particular of *prospective* longitudinal panel studies is recommended.



In this regard, one major shortcoming of many of the prospective panel studies described above relates to the fact that several key aspects of internal migration are not represented. It is recommended that datasets should at least include information on the place of residence (at the smallest possible spatial level), typologies of the characteristics of the place of residence, information on a change of residence, reasons for moving, the plan to migrate, information on the dwelling and the neighborhood, as well as on commuting and multi-local living arrangements. For the purpose of cross-national comparisons, information on the place of residence should be made available in form of the EU's Nomenclature of Territorial Units for Statistics, or NUTS (*Nomenclature des unités territoriales statistiques*) where the level three NUTS code corresponds to rural or city districts (*ländliche Kreise* and *kreisfreie Städte*).

If the structural characteristics of the residential environment are not included, information on the location of individual residences should at least be combinable with spatial context information from other aggregated regional datasets. Particularly for the analysis of intra-urban moves, regional context information must be provided on a very small scale. Matching survey data with geographical microdata from MICROM is a significant step forward. Consideration should also be given to matching survey data with small-scale spatial data from the Inner City Spatial Monitoring (IRB, *Innerstädtische Raumbeobachtung*) of the BBR. Moreover, the typology of inner-city location types (*innerstädtische Lagetypen*) used by Inner City Spatial Monitoring should be implemented in the datasets.

Since the decision to migrate is very complex, further opportunities to analyze this process by using a multi-actor design should be provided. This implies accounting for structural characteristics, attitudes, and the decisions of other individuals in the person's household or even in their larger social network.

The most comprehensive longitudinal dataset is the SOEP, which collects structural and non-structural information on the dynamics of housing conditions and residential moves. Still, the value of this dataset is restricted by the general fact that residential moves do not occur that often during the lifetime. It thus follows that for some research issues, notably analyses of specific migration types (e.g., intra-urban moves), the size of the (sub)-sample becomes too small and therefore is no longer representative. One solution of this limitation could be to increase the sample size of the SOEP.

In general, regional multi-stage cluster sampling techniques should be used to collect data for internal migration research to assure regional type-specific analyses. A possible typology especially for inner-city cluster sampling could be the inner-city location types used by the Inner City Spatial Monitoring of the BBR.

## References:

- Arbeitsgruppe Regionale Standards (2005): Regionale Standards – Eine gemeinsame Empfehlung des Arbeitskreises Deutscher Markt- und Sozialforschungsinstitute e.V. (ADM), der Arbeitsgemeinschaft Sozialwissenschaftlicher Institute e.V. (ASI) und des Statistischen Bundesamtes, ZUMA. Mannheim.
- Birg, H./Flöthmann, E.J./Heins, F. and Reiter, I. (1993): Migrationsanalyse. Forschungen zur Raumentwicklung, vol. 22. Bonn.
- Bleck, M. and Wagner, M. (2006): Stadt-Umland-Wanderung in Nordrhein-Westfalen – eine Meta-Analyse. *Raumforschung und Raumordnung* 2/2006, 104-115.
- Borjas, G.J. (1987): Self-selection and the earnings of immigrants. *American Economic Review* 77 (4), 531-553.
- Börsch-Supan, A./Jürges, H. and Lipps, O. (2003): SHARE – Measuring the ageing process in Europe. *Zuma Nachrichten* 53, 96-113.
- Brühl, H./Echter, C.P./Frölich von Bodelschwingh, F. and Jekel, G. (2005): Wohnen in der Innenstadt – eine Renaissance? *Difu-Beiträge zur Stadtforschung*. Berlin.
- Büchel, F. and Schwarze, J. (1994): Die Migration von Ost- nach Westdeutschland – Absicht und Realität. *Mitteilungen aus der Arbeitsmarkt- und Berufsforschung* 27 (1), 43-52.
- Bundesamt für Bauwesen und Raumordnung (Ed.) (2007): Innerstädtische Raumbewachung: Methoden und Analysen. Bonn.
- Burda, M.C. (1993): The determinants of East-West German migration: Some first results. *European Economic Review* 37 (2-3), 452-461.
- De Jong, G.F. and Fawcett, J.T. (1981): Motivations for migration: an assessment and a value-expectancy research model. In: De Jong, G.F. and Gardner, R.W. (Eds.): *migration decision making*. New York.
- Drews, N. (2008): Das Regionalfiler der IAB-Beschäftigtenstichprobe 1975-2004. Handbuch-Version 1.0.0, FDZ Datenreport. Nuremberg.
- Eichstädt-Bohlig, F./Hannemann, Ch. and Bohne, R. (2006): Die Gestaltung der schrumpfenden Stadt. In: Heinrich-Böll-Stiftung (Ed.): *Das neue Gesicht der Stadt*. Berlin.
- Farwick, A. (2001): Segregierte Armut in der Stadt. Ursachen und soziale Folgen der räumlichen Konzentration von Sozialhilfeempfängern. Opladen.
- Florida, R. (2004): *The rise of the creative class*. New York.
- Friedrich, K. (1995): Altern in räumlicher Umwelt. Sozialräumliche Interaktionsmuster älterer Menschen in Deutschland und den USA. Darmstadt.
- Friedrich, K. and Schultz, A. (Eds.) (2008): *Brain drain oder brain circulation? Konsequenzen und Perspektiven der Ost West Migration*, Leibnitz-Institut für Länderkunde. Leipzig.
- Friedrichs, J./Häußermann, H. and Siebel, W. (Eds.) (1986): *Süd-Nord Gefälle in der Bundesrepublik? Opladen/Wiesbaden*.
- Gatzweiler, H.P. (1975): Zur Selektivität interregionaler Wanderung. Ein theoretisch-empirischer Beitrag zur Analyse und Prognose altersspezifischer interregionaler Wanderung, vol. 1. Bonn.
- Goebel J./Spieß C.K./Witte, N. and Gerstenberg, S. (2007): Die Verknüpfung des SOEP mit MICROM-Indikatoren: Der MICROM-SOEP Datensatz. DIW Data Documentation 26.

- Heckman, J.J. (1979): Sample selection bias as a specification error. *Econometrica* 47, 153-161.
- Hoffmeyer-Zlotnik, J.H.P. (2005): Regionalisierung sozialwissenschaftlicher Umfragedaten – Eine Einführung. In: Grötzinger, G. and Matiaske, W. (Eds.): *Deutschland regional*. Munich, 3-28.
- Huinink, J. and Kley, S. (2008): Regionaler Kontext und Migrationsentscheidungen im Lebensverlauf. In: Kalter, F. (Ed.): *Migration, Integration und ethnische Grenzen*. Sonderband der Kölner Zeitschrift für Soziologie, 162-184.
- Hunt, J. (2004): Are migrants more skilled than non-migrants? Repeat, return, and same-employer migrants. *Canadian Journal of Economics* 37 (4), 830-849.
- Jürges, H. (1998): Beruflich bedingte Umzüge von Doppelverdienern. Eine empirische Analyse mit Daten des SOEP. *Zeitschrift für Soziologie* 27, 358-377.
- Kalter, F. (1994): Pendeln statt Migration? Die Wahl und Stabilität von Wohnort-Arbeitsort-Kombinationen. *Zeitschrift für Soziologie* 23, 460-476.
- Kalter, F. (1997): Wohnortwechsel in Deutschland. Ein Beitrag zur Migrationstheorie und zur empirischen Anwendung von rational-choice Modellen. Opladen.
- Lee, E.S. (1972): A theory of migration. *Demography*, 45-57.
- Lüttinger, P. and Riede, Th. (1997): Der Mikrozensus: amtliche Daten für die Sozialforschung. *ZUMA-Nachrichten* 41, 19-43.
- Massey, D.S. and Espinosa, K.E. (1997): What's driving Mexico-U.S. migration? A theoretical, empirical, and policy analysis. *The American Journal of Sociology* 102 (4), 939-999.
- Ravenstein, E.G. (1995): The Laws of Migration. [Reprint of the title originally publ. in 2 parts in vol. 48, pt. 2, June 1885 and vol. 52, pt. 2, June 1889 of the *Journal of the Royal Statistical Society*. North Stratford].
- Rendtel, U. (2005): Wie geeignet ist der Mikrozensus als Datenbasis für Längsschnittanalysen? Methodenverbund „Aufbereitung und Bereitstellung des Mikrozensus als Panelstichprobe“. Arbeitspapier Nr. 7. Wiesbaden.
- Rossi, P.H. (1955): *Why families move: A study in the social psychology of urban residential mobility*. Illinois.
- Ruckdeschel, K./Ette, A./Hullen, G. and Leven, I. (2006): *Generations and Gender Survey. Dokumentation der ersten Welle der Hauptbefragung in Deutschland. Materialien zur Bevölkerungswissenschaft*. Wiesbaden.
- Rusconi, A. (2006): *Leaving the parental home in Italy and West Germany: Opportunities and constraints*. Aachen.
- Schlömer, C. and Bucher, H. (2001): Arbeitslosigkeit und Binnenwanderungen. Auf der Suche nach einem theoriegestützten Zusammenhang. *Informationen zur Raumentwicklung* 1, 33-47.
- Sjastaad, L.A. (1962): The Costs and Returns of Human Migration. *The Journal of Political Economy* 70, 80-93.
- Spieß, C.K. (2005): Das Sozio-oekonomische Panel (SOEP) und die Möglichkeiten regionalbezogener Analysen. In: Grötzinger, G. and Matiaske, W. (Eds.): *Deutschland regional*. Munich.
- Todaro, M.P. (1969): A model of labor migration and urban unemployment in less developed countries. *American Economic Review* 59, 138-48.
- Wagner, G.G. (1992): Arbeitslosigkeit, Abwanderung und Pendeln von Arbeitskräften der neuen Bundesländer. *Sozialer Fortschritt* 4, 84-89.

- Wagner, G.G./Frick, J.R. and Schupp, J. (2007): The German Socio-Economic Panel Study (SOEP) – Scope, Evolution and Enhancements. *Schmollers Jahrbuch* 127 (1), 139-169.
- Wagner, M. (1989): *Räumliche Mobilität im Lebensverlauf*. Stuttgart.
- Windzio, M. (2004): Zwischen Nord- und Süddeutschland: Die Überwindung räumlicher Distanzen bei der Arbeitsmarktmobilität. *Zeitschrift für Arbeitsmarktforschung* 1, 29-44.
- Windzio, M. (2007): Regionale Arbeitslosigkeit und Distanz zur Grenze: Individual- und Kontexteffekte auf die Abwanderung von Arbeitskräften von Ost- nach Westdeutschland. *Schmollers Jahrbuch* 127 (4), 553-583.
- Windzio, M. (2009): The “Exit Option” of Labour Migration from East to West-Germany: Individual and Contextual Determinants of Geographic Mobility of Unemployed Workers. In: Egbert, H. and Kolb, H. (Eds.): *Migrants and Markets: Perspectives from Economics and the Other Social Sciences*. Amsterdam.
- Wirth, H./Zühlke, S. and Christians, H. (2005): Der Mikrozensus als Datenbasis für die Regionalforschung. In: Grötzing, G. and Matiaske, W. (Eds.): *Deutschland regional*. Munich.
- Wolpert, J. (1965): Behavioral Aspects of the Decision to Migrate. *Papers and Proceedings of the Regional Science Association* 15, 159-169.

## 1.4 Fertility and Mortality Data for Germany

*Michaela Kreyenfeld and Rembrandt Scholz\**

---

\* We would like to thank Klaus Duschek for his detailed and friendly responses to our various inquiries about the Microcensus. For language editing, we would like to thank Miriam Hils.

Contact:

Michaela Kreyenfeld  
Rembrandt Scholz  
Max Planck Institute for Demographic Research  
Konrad-Zuse-Str. 1  
18057 Rostock  
Germany  
e-mail: [kreyenfeld\[at\]demogr.mpg.de](mailto:kreyenfeld[at]demogr.mpg.de)  
[scholz\[at\]demogr.mpg.de](mailto:scholz[at]demogr.mpg.de)

## Abstract

There has been considerable progress made in the improvement of the data infrastructure for fertility and mortality researchers in Germany in recent years. Several large-scale datasets have been made available through the establishment of Research Data Centers. The Microcensuses of the 1970s and 1980s, the censuses of the German Democratic Republic and the Federal Republic of Germany, the Microcensus panel, data from the pension registers, individual-level data from vital statistics, and the central foreigner registers are now available for scientific use. Vital statistics have been reformed, and the Microcensus now collects information on the number of children a woman has given birth to during her life. Despite these improvements, there are still some “weak spots” in Germany’s data infrastructure. Germany is lacking official counts of reconstituted families. We know little about the mortality risks of immigrants. In addition, the data infrastructure for studying socio-economic differences in mortality risks could be improved, enabling Germany to catch up with international developments in this area. This paper concludes by making some suggestions for improving the available data.

Keywords: fertility, mortality, demographic data

## 1. Introduction

Since the turn of the last century, demographic change has been a popular topic among journalists and policy-makers alike. Yet despite the considerable level of public interest in this topic, the available data was rather poor in Germany: important fertility indicators were lacking, official mortality rates for the “oldest old” were of poor quality, and population counts were inaccurate. Today, however, we can state that the data situation for researchers interested in the field of demographic change has improved tremendously. Germany is about to conduct a register-based census which is expected to give an accurate account of the population size in Germany. Furthermore, new micro-level datasets have become available for scientific use that will enhance our understanding of demographic processes.

This paper provides an overview of what we believe are the most important recent innovations in the field of fertility and mortality research. Obviously, such an overview is subjective and can not be considered comprehensive. Nevertheless, we believe that we have included the most significant and critical datasets in this brief overview. Part 2 presents data and discusses applications. In Part 3 we discuss what could be done to improve data availability in the future. Part 4 concludes the overview, and provides a list of recommendations for the future.

## 2. Recent progress in the data infrastructure

### 2.1 Fertility and family research

In the field of family and fertility research, an important step forward was made recently with the amendment of the German Population Statistics Law (*Bevölkerungssstatistik-Gesetz*), which prescribes which data are to be collected for population statistics. For centuries, German vital statistics did not collect births by biological order. Since 2008, German vital statistics includes this type of information (Deutscher Bundestag 2007). Another important amendment provides that the Microcensus will ask female respondents to give the number of their biological children.<sup>1</sup>

#### Age at first birth and childlessness

The groundbreaking changes in the law will enable researchers to generate important structural fertility indicators, such as the mean age at first birth. The postponement of first birth is one of the most important changes in fertility behavior of the recent years (Sobotka 2004; Billari et al. 2006). Germany has been a forerunner in this development, but official indicators documenting this process were lacking. Due to the amendment of the German Population Statistics Law, it is now possible to generate a (period) mean age at first birth. This indicator is of great public interest. Furthermore, it is a measure that will enter international demographic statistics.

In addition to changes in the age at first birth, the level of childlessness is an indicator that is in great demand and frequently discussed (Berth 2005; Mönch 2007; Schwentker 2007). However, the ultimate level of childlessness cannot yet be calculated based on German vital statistics.<sup>2</sup> This gap in vital statistics can be filled through other sources, however. The *Frauenbefragung Geburten* has been an important source of indicators of permanent childlessness (Pötzsch 2007). In the future, the Microcensus will provide this information, too.

---

1 The plan is to collect this information every four years. The Microcensus 2008 is the first to include a question on whether the respondent has ever given birth to a child and another one on the total number of children ever born. The question will be asked of female respondents aged 15 to 75.

2 Since 2008, German vital statistics provide birth order-specific fertility information, which is needed to calculate indicators of childlessness. However, birth order specific fertility information for the entire reproductive life of a cohort must be collected first before an ultimate level of childlessness can be generated. The cohort 1993 will be the first one for which birth order information will be available for their entire reproductive lives. This cohort reaches the end of their fertile years in 2038.



### Fertility of migrants

From 2008 onwards, the Microcensus will enable researchers to generate fertility indicators according to the socio-economic characteristics of the respondents. This will also enable us to generate the total number of children by nationality and migration background. In addition to the Microcensus, the Turkish sample of the Generations and Gender Survey (GGS) will complement our understanding of the demographic behavior of non-citizens and migrants. The fertility of migrants is an aspect worth pointing out, not only because this topic is of great scientific interest (see Nauck 2007; Milewski 2007), but also because vital statistics are not very useful for understanding the fertility dynamics of non-citizens and migrants. This is partly because population counts of non-German citizens have been imprecise. But this also relates to the fact that it is difficult to generate fertility indicators for a highly mobile population with aggregate level data.

### Panel studies in the field of family and fertility

In the past, the Socio-Economic Panel (SOEP, *Sozio-oekonomisches Panel*) has been the major panel study for family and fertility researchers. Although this data provides a rich battery of socio-economic variables, it does not provide much information on the quality of partnership or the intention to become parent. This has limited our opportunities to study, for example, how fertility intentions transfer into behavior. Germany now provides two important datasets – the GGS and the Panel Analysis of Intimate Relationships and Family Dynamics (pairfam) – that will help shed light on the decision-making processes that underlie fertility and nuptiality behaviors. The first wave of the GGS has been released (Ruckdeschel et al. 2006). Data from the second round of the GGS, as well as data of the pairfam, were collected in autumn 2008 (Feldhaus and Huinink 2008).

### Fertility and large-scale datasets

For demographic studies, having access to large-scale data is indispensable. In this context, the great achievement of the Research Data Centers must be acknowledged. The Research Data Centers of the Federal Statistical Office and the Statistical Offices of the German *Länder* have made available individual-level data for births and marriages. Additionally, the Microcensuses of the 1970s and 1980s and the censuses of the German Democratic Republic and the Federal Republic of Germany have been made available for scientific purposes. The Scientific Use File of the Microcensus, which opens up new potential for fertility and family analysis, is also now accessible (Schmidtke et al. 2008). Finally, the Research Data Center of the German

Pension Insurance (RV, *Deutsche Rentenversicherung*) provides researchers access to Scientific Use Files of pension records, which can also be used for fertility and family research (in particular, the *Versicherungskontenstichprobe*) (Kreyenfeld and Mika 2008). Fertility analyses with register data, like those that have previously been undertaken mainly for Scandinavian countries, can now be replicated using German data.

## 2.2 Aging and mortality

It is as crucial in the field of mortality and aging to have access to large-scale datasets as it is for demographic studies. After all, death is quite a rare event. Therefore, large datasets are needed in the calculation of robust mortality estimates. Fortunately, there has been considerable progress made in recent years in terms of the availability of large-scale datasets. New computer techniques and opportunities offered by installed process data sources enable researchers to conduct mortality analyses on large-scale datasets.

### Human Mortality Database, population size and census

The Human Mortality Database (HMD) is a collaborative project which has been conducted since 2002 by the University of California at Berkeley (US) and the Max Planck Institute for Demographic Research (Rostock, Germany). The purpose of the database is to provide researchers around the world with easy access to detailed and comparable national mortality data via the Internet.<sup>3</sup> The HMD methodology has been used to validate German population statistics. In Germany, the last census was conducted in the West in 1987, and in the former East Germany in 1990.<sup>4</sup> In order to generate the population size, German vital statistics largely relies on the results from the last census and a component-method by births, immigrations, out-migrations, and deaths (*Bevölkerungsfortschreibung*). There is reason to believe that the population estimates that are generated from the *Bevölkerungsfortschreibung* are particularly distorted with growing distance to the last census, especially among the highest ages.

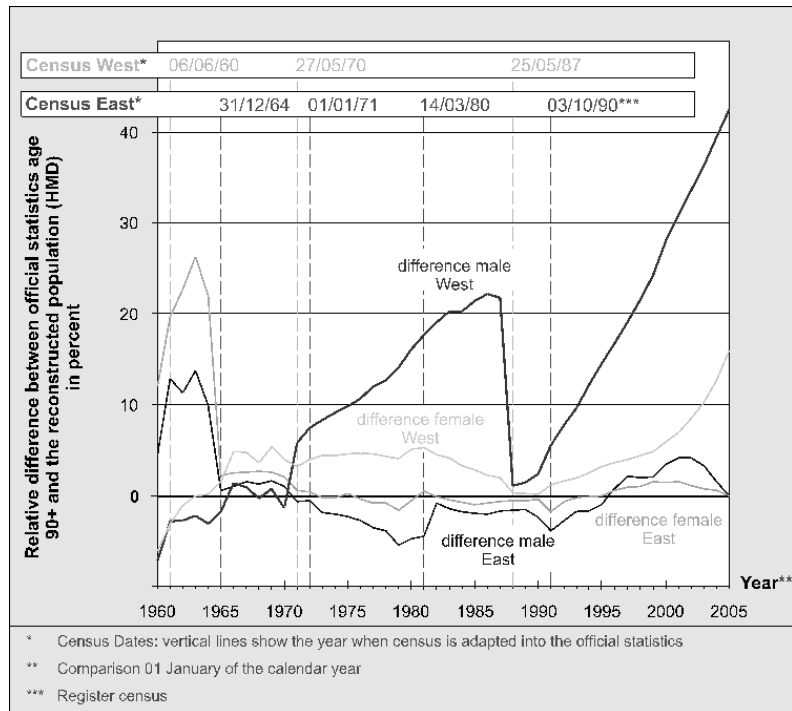
---

3 The database will contain original life tables for thirty-five countries, as well as all raw data used in constructing those tables. The raw data generally consist of birth and death counts from vital statistics, population counts from periodic censuses, and official population estimates. The general documentation and the steps followed in computing mortality rates are described in the methods protocol. There are datasets for East Germany, West Germany, and Germany Total for the period of 1955-2007 (<http://www.mortality.org/>).

4 The last "conventional" census of the German Democratic Republic was conducted in 1981. However, there were population registers in East Germany which provide reliable population counts. These registers were discontinued in 1990.

As shown in figure 1, the difference between the official and the recalculated population for men age 90+ in West Germany grows with the amount of time that has passed since the last census. Just after the West German Census in 1987, a sudden jump can be seen in the official population. This suggests that the population of males age 90 and over is strongly overestimated in German vital statistics (for more detail, see Jdanov et al. 2005; Scholz and Jdanov 2006). It may be hoped that the new census will improve the quality of the data available for studying mortality at higher ages.

Figure 1: Comparison of relative differences in population estimates in Germany 1960-2005 between official statistics and HMD



Source: own estimations

### Socio-economic inequality in old age mortality

Relative socio-economic inequality in old age mortality is a major public health issue given the growing size of the elderly population and the sharp rise in absolute mortality with age. In the past, the international literature in this area was marked by the persistent absence of Germany. In many reviews of socio-economic mortality differences in Europe, Germany was not included. One reason for this is that, unlike in many other countries, German population statistics do not provide suitable data for mortality estimation by socio-economic status. Social science surveys can only partially fill this gap since the number of elderly subjects is too small for a robust estimation of mortality differentials in this kind of data. Furthermore, the survey data suffer from recruitment bias and the absence of people living in institutions. However, the situation has changed recently with the introduction of new policies enabling scientific analyses of administrative microdata. Data from the Research Data Center of the German Pension Insurance can now be used to evaluate mortality differentials among men aged 65 and older (Gaudecker and Scholz 2007; Shkolnikov et al. 2008; Himmelreich et al. 2008).

### The healthy migrant effect

It is known from several studies that migrants are healthier, and thus show lower mortality than the native population. This phenomenon has been described for various countries and ethnic groups, and holds true for both internal and international migrants. Generally, this development is explained by a special selection effect which may influence mortality and morbidity rates. This selective migration is thought to operate in two directions, which involve the movement of a “select group” of either the healthy or the unhealthy. The movement of healthier individuals is known as the so-called “healthy migrant effect.” Conversely, it appears that sick migrants are involved in return migration, in order, for example, to be closer to family or to care-giving institutions. The latter phenomenon is also known as “salmon bias.” For Germany, it is also important to consider whether migrants’ low mortality rates are caused by inaccuracies in the vital statistics, for example, if doubtful data quality might contribute to migrants’ “statistical immortality.” Newly available data will help shed more light on this issue, specifically the Immigrant Survey of the Federal Institute for Population Research (BiB, *Bundesinstitut für Bevölkerungsforschung*) in Wiesbaden (Luy 2007), data from the German Pension Insurance (Kibele et al. 2008), and data from the German Central Register of Aliens (AZR, *Ausländer-zentralregister*) (Kohls 2008).

### 3. Challenges and recommendations

Overall, the infrastructure for conducting fertility and mortality research has improved tremendously in recent years. Nevertheless, there are some “weak spots” in Germany’s data infrastructure, which we will discuss in the following.

#### Family change and official statistics

Official statistics have always been slow in catching up to changes in the family. For a long time, the official UN definition of what is a family ignored new family forms, such as non-marital unions with children. This has changed in the recent years. In the UN recommendations for what is to be included in the census, co-residential partnerships are named among the core concerns (UN 2006: 113). Germany will be able to provide counts on co-residential unions based on data from the Microcensus. One drawback is that the question about partnership status, which is needed to identify a non-marital union, is voluntary, and about five percent of respondents refuse to answer the question (Heidenreich and Nöthen 2002). Since the share of non-marital unions has become such an integral demographic indicator, it seems odd that partnership status is one of the few questions in the Microcensus for which a response is not compulsory.

A related issue concerns stepfamilies. Families in which children live with biological and/or non-biological parents are on the rise, and they pose important new social policy questions. However, we do not have an accurate account of the share of reconstituted families in Germany. In the census, more complex living arrangements, such as stepfamilies, cannot be identified – despite the fact that the UN (2006) requested that this information be included in the census. Survey data, such as data from the Generations and Gender Survey, provide detailed information on family structure and living arrangements. However, sample sizes are too small to provide good “structural indicators” on the prevalence of reconstituted families. In the Microcensus, it is difficult to identify “stepfamily constellations,” because the kinship status of the household members is only surveyed in reference to the head of the household.

It is difficult to make recommendations for resolving this problem. The household relationship matrix is usually seen as a superior method for surveying living arrangements (Statistical Commission and UN Economic Commission for Europe / Statistical Office of the European Communities; UN 2006: 107). If this method were introduced into the Microcensus, the share of stepfamilies in Germany could be established. However, this would obviously require a fundamental change in the Microcensus questionnaire.

Another solution could be to find out whether respondents may be asked if the stepparent, adoptive parent, or foster mother or father lives in the same household.<sup>5</sup>

#### Piecemeal changes in the field of family and fertility

While there has been significant progress made in improving Germany's data infrastructure, some changes remain incomplete. For example, it is certainly a great achievement that the number of biological children is now counted in the German Microcensus. However, it seems unfortunate that only women are asked about their fertility careers, since male fertility is also an important area for fertility and family researchers (Tölke and Hank 2005). In the social science dataset, it has become standard to ask both female and male respondents about their fertility careers. It seems socially regressive that, in the Microcensus, males have been filtered around the question concerning the number of biological children.

Finally, the Microcensus for fertility research would be tremendously enhanced if it included information on the ages at first, second, and additional births. Such a suggestion would certainly provoke another heated debate about whether the Microcensus questionnaire is already overloaded. However, a simpler solution could be found by repeating the *Frauenbefragung Geburten* on a regular basis to provide structural indicators of fertility change in Germany.

#### Socio-economic differences in mortality risks

In the field of mortality research, we must conclude that we still know too little about the mortality risks of immigrant populations. The data infrastructure for studying socio-economic differences in mortality risk could also be improved to keep pace with international developments in this area. We simply know too little about how mortality risks differ in Germany by educational level and socio-economic status. One way to improve this situation could be to establish a central mortality register similar to those that exist in other countries, such as Sweden or the US (Mueller 2008). However, this type of initiative will surely have to pass several administrative hurdles. An easier solution may be found by investigating ways that the Microcensus panel could be used for mortality research. Currently, it cannot be used for

---

5 The Microcensus already includes a question on whether the mother or father lives in the same household. However, it does not allow the respondent to distinguish whether he or she is a stepparent, adoptive parent, or foster parent. Legal regulations might make it impossible to ask respondents whether they have adoptive parents. However, a distinction between foster parents, stepparents, and biological or adoptive parents might present less of a legal problem.

this purpose because there is no systematic documentation of information on respondents who drop out – whether because they die or move to a different location. Finding a way to collect this information would not only increase the potential for using the Microcensus panel in mortality research, but would also expand the possibilities for employing the Microcensus panel in many other kinds of longitudinal research.

#### **4. Conclusion**

In this paper, we have described the significant progress that has been made in improving the data infrastructure for research on fertility and mortality in Germany. Nevertheless, additional changes and improvements could be made that would further increase our understanding of fertility and mortality processes. In terms of furthering research in these areas, we have argued that we need better structural indicators to capture family change in Germany. This would include making official counts of *reconstituted families* and also raises the possibility of making the question on *partnership status* compulsory in the Microcensus. In the field of mortality research, we stressed that we need better estimates of the *mortality risks of migrants*, and a better understanding of the *socio-economic determinants of death*. In this context, we pointed out the potential of the Microcensus to help fill the gap in data collection.

## References:

- Berth, F. (2005): Die demografischen Märchen: eine neue Gebärstreik-Debatte in Deutschland. *Süddeutsche Zeitung*: November 8, 2005.
- Billari, F.C./Liefbroer, A.C. and Philipov, D. (2006): The postponement of childbearing in Europe: Driving forces and implications. *Vienna Yearbook of Population Research* 2006, 1-17.
- Deutscher Bundestag (2007): Entwurf eines Gesetzes zur Änderung des Mikrozensusgesetzes 2005 und des Bevölkerungsstatistikgesetzes. Drucksache 16/5239.
- Feldhaus, M. and Huinink, J. (2008): Neuere Entwicklungen in der Beziehungs- und Familienforschung. *Schriften zum Beziehungs- und Familienentwicklungspanel*. Band 1. Würzburg.
- Fishbein, M. and Ajzen, I. (1980): *Understanding attitudes and predicting social behavior*. Upper Saddle River.
- Gaudecker, H.-M. v. and Scholz, R.D. (2007): Differential mortality by lifetime earnings in Germany. *Demographic Research* 17 (4), 83-108.
- Heidenreich, H.-J. and Nöthen, M. (2002): Der Wandel der Lebensformen im Spiegel des Mikrozensus. *Wirtschaft und Statistik* 1/2002, 26-38.
- Himmelreicher, R./Sewöster, D./Scholz, R.D. and Schulz, A. (2008): Die fernere Lebenserwartung von Rentnern und Pensionären im Vergleich. *WSI-Mitteilungen* 5/2008, 274-280.
- Jdanov, D.A./Scholz, R.D./Shkolnikov, V.M. (2005): Official population statistics and the Human Mortality Database estimates of populations aged 80+ in Germany and nine other European countries. *Demographic Research* 13, 335-362.
- Kibele, E./Scholz, R.D. and Shkolnikov, V.M. (2008): Low migrant mortality in Germany for men aged 65 and older: fact or artifact? *European Journal of Epidemiology* 23 (6), 389-393.
- Kohls, M. (2008): *Leben Migranten wirklich länger?* Bundesamt für Migration und Flüchtlinge. Working Paper der Forschungsgruppe des Bundesamtes / Bundesamt für Migration und Flüchtlinge No. 16.
- Kreyenfeld, M. and Mika, T. (2008): *Erwerbstätigkeit und Fertilität: Analysen mit der Versicherungskontenstichprobe der deutschen Rentenversicherung*. Deutsche Rentenversicherung (Sonderausgabe) Band 79, 71-95.
- Luy, M. (2007): Estimating the migrant survival advantage from orphanhood of second generation migrants. *Rostocker Zentrum – Diskussionspapier* No. 17 (November 2007).
- Milewski, N. (2007): First child of immigrant workers and their descendants in West Germany: Interrelation of events, disruption, or adaptation? *Demographic Research* 17, 859-896.
- Mönch, R. (2007): *Frau Doktors unsichtbares Kind*. *Frankfurter Allgemeine Sonntagszeitung*: June 10, 2007
- Mueller, U. (2008): *Argumente für die Einrichtung eines nationalen Mortalitätsregisters*. RatSWD Working Paper No. 40.
- Nauck, B. (2007): Immigrant families in Germany. Family change between situational adaptation, acculturation, segregation and remigration. *Zeitschrift für Familienforschung* 19 (1), 34-54.



- Pöttsch, O. (2007): Neue Datenquelle zu Geburten und Kinderlosigkeit. *Wirtschaft und Statistik* 3/2007, 260-263.
- Ruckdeschel, K./Ette, A./Hullen, G. and Leven, I. (2006): Generations and Gender Survey. Dokumentation der ersten Welle der Hauptbefragung Deutschland. Wiesbaden.
- Schmidtke, K./Kreyenfeld, M. and Zühlke, S. (2008): Ökonomische Basis der Familiengründung – Analysen mit den Daten des Mikrozensus-Panels. [Paper presented at the „Statistische Woche“, Cologne 2008].
- Scholz, R.D. and Jdanov, D.A. (2006): The use of the data of the Forschungsdatenzentrum (FDZ) of the pension system in Germany – A correction method for population estimates in old age. In: DRV (Ed.): Bericht vom dritten Workshop des FDZ der Rentenversicherung (FDZ-RV). DRV-Schriften Band 55/2006, 200-211.
- Schwentker, B. (2007): Ende einer Diskriminierung. *ZEIT online*: July 6, 2007.
- Shkolnikov, V.M./Scholz, R.D./Jdanov, D.A./Stegmann, M. and Gaudecker, H.M. v. (2008): Length of life and the pensions of five million retired German men. *The European Journal of Public Health* 18 (3), 264-269.
- Sobotka, T. (2004): Is Lowest-Low Fertility in Europe Explained by the Postponement of Childbearing? *Population and Development Review* 30 (2), 195-220.
- Statistical Commission and UN Economic Commission for Europe/Statistical Office of the European Communities (2004): Towards possible changes to the census recommendations on families and household. Geneva.
- Tölke, A. and Hank, K. (2005): Männer – Das "vernachlässigte" Geschlecht in der Familienforschung. Wiesbaden.
- United Nations (UN) (2006): Conference of European Statisticians Recommendations for the 2010 Censuses of Population and Housing. Geneva



## 2. MEASURING COMPETENCIES

### 2.1 Measuring Cognitive Ability

*Roland H. Grabner and Elsbeth Stern*

Contact:

Roland H. Grabner  
Elsbeth Stern  
Swiss Federal Institute of Technology (ETH) Zurich  
Institute for Behavioral Sciences  
Universitätsstrasse 6  
8092 Zurich  
Switzerland  
e-mail: grabner[at]ifv.gess.ethz.ch  
stern[at]ifv.gess.ethz.ch

## **Abstract**

The assessment of cognitive abilities is critical in large-scale survey studies that aim at elucidating the longitudinal interplay between the individual's cognitive potential and socio-economic variables. The format of such studies calls for assessment methods that not only can be efficiently administered, but also show a high level of (psychometric) measurement quality. In consideration of recent theoretical and empirical advances in intelligence research, we recommend the implementation of tests drawing on working memory in large-scale survey studies. Working memory is a limited-capacity system for the temporary storage and processing of information that is currently considered to be the key cognitive system underlying intellectual abilities. Examples of four types of working memory tests are described and critically evaluated with regard to their psychometric quality and the need for further evaluation.

**Keywords:** cognitive abilities, intelligence, knowledge, information processing, mental speed, working memory

## **1. Research questions and theoretical developments**

The analyses of gene-environment interaction and evolution are becoming more and more accepted as a research focus in the social sciences (see Spinath 2008). The basic argument is that without the “control” of genetic effects one cannot be sure that he or she is estimating unbiased socio-economic effects (Guo 2008; Diewald 2008). In this context, cognitive abilities play an important role. Cognitive abilities are the raw material for developing individual resources and are both promoted as well as constrained by socio-economic context. Research on cognitive abilities has revealed considerable associations between an individual's cognitive abilities (“general intelligence”) and numerous indicators of life success – ranging from educational and vocational performance to delinquency, morbidity, and mortality (Jensen 1998; Deary et al. 2004). The causal nature of most of these correlations is still unknown, as is the mediating role of socio-economic variables. This underlines the importance of including cognitive ability measurements in large-scale survey studies to enhance our knowledge about the longitudinal interplay between individual cognitive resources and socio-economic variables.

### *1.1 A brief history of cognitive ability assessment*

The first systematic approach to objectively measuring cognitive abilities can be traced back to Sir Francis Galton at the end of the nineteenth century (e.g., Jensen 1998). Galton developed a variety of tests to measure elementary mental functions such as sensory discrimination and perception speed, guided by the assumption that differences in intellectual ability result from a differential efficiency of the central nervous system. Galton's tests were presented to the public in his Anthropometric Laboratory at the International Health Exhibition in London. The interest into the new anthropometric measurements was enormous; between 1884 and 1890, the data of more than 9000 persons were collected. The validity of the tests for measuring cognitive abilities, however, was disappointing. The test results turned out to be only poorly correlated with commonsense criteria of intellectual abilities and educational success.

A more promising approach was pursued by Alfred Binet at the beginning of the twentieth century (1905). Commissioned by the French Ministry of Public Instruction, Binet was charged with developing a quick and reliable method of distinguishing mentally retarded children – who were not expected to profit from normal instruction in school – from those with mere behavior problems. In contrast to Galton, Binet and his colleague Simon devised a battery of tasks drawing on practical knowledge and skills rather than on elementary mental functions. Children were instructed to point at various parts of their body, name objects seen in a picture, give definitions, repeat series of digits or a complete sentence, tell the time of a clock, etc. Besides their focus on relatively practical skills, Binet and Simon's approach was innovative in that they used the children's age as an external criterion for cognitive abilities. By empirically assigning the tasks to different age groups, their intelligence scales allowed the objective assessment of whether a child was advanced or backward for his or her chronological age and, thus, to distinguish mentally retarded children from others. This comparison of mental with chronological age provided the basis for the advent of the intelligence quotient (IQ) (Stern 1912), until it was replaced by the concept of today's statistical deviation IQ (Wechsler 1944).

The Binet scales were soon translated and distributed in America and England and became the norm against which later intelligence tests were evaluated. The further development of intelligence tests was strongly related to the question of the structure of cognitive abilities. At a gross level, two different views can be distinguished. Some researchers (e.g., Jensen) emphasized the existence and importance of a general intelligence (*g*) factor, which was originally discovered by Spearman (1904). If a large and random sample of participants completes a number of diverse cognitive tests, the correlations among the different test scores will be almost entirely positive

and, in most of the cases, of moderate size. This means that a person who does well in one test also has a high probability of achieving a good performance level in the other tests. Using statistical methods, this correlation pattern can be reduced to one single factor (the *g* factor), which usually accounts for about 50 percent of the entire test variance.

Other researchers (e.g., Thurstone or Gardner), in contrast, questioned the existence of the *g* factor. This diverging view predominantly resulted from the application of different statistical methods in analyzing cognitive test performance data or from the expansion of the intelligence concept to non-academic skills (such as interpersonal and bodily-kinaesthetic intelligence; Gardner 1983).

At present, there is a wide consensus on a hierarchical model of cognitive abilities, consisting of three levels of different generality (Carroll 1993; Gustafsson 1984). At the top and most general level is Spearman's *g* factor, reflecting the fact that diverse cognitive abilities show near-universal positive correlations.

Group factors for cognitive abilities such as fluid and crystallized intelligence are located at the second level. Fluid intelligence is conceptualized as the ability to solve novel problems and is typically assessed by tasks drawing on abstract reasoning (inductive or deductive) or complex problem solving. Crystallized intelligence reflects the breadth and depth of general knowledge and is usually measured by tests on vocabulary, spelling ability, or general information.

Finally, at the lowest level, there are specific cognitive abilities such as quantitative reasoning (for fluid intelligence) or lexical knowledge (for crystallized intelligence), accounting for variance that is neither attributable to factor *g* nor to the group factors. Although hierarchical models with *g* at the top and second- and third-order factors below might best describe the structure of individual differences in cognitive abilities, it is also widely accepted that most of the predictive value of intelligence tests derives from the *g* factor, which is strongly related to fluid intelligence (Brody 1999; Deary 1998; Jensen 1998).

## *1.2 Bases of cognitive abilities*

Over the past decades, a great deal of research has been conducted to better understand the bases of individual differences in cognitive abilities. At present, two cognitive components are discussed that show consistent associations with intelligence and might, therefore, be considered as potential bases of human intelligence. The first component is *mental speed* (e.g., Neubauer 1995). There is a large body of evidence showing consistent negative associations between intelligence and reaction times in so-called elementary cognitive tasks (ECTs). ECTs are designed to place only minimal require-

ments on the participant and, thus, are less likely to be influenced by differential strategies or prior knowledge. As an example, in the letter-matching task by Posner and Mitchell (1967), the participants have to judge whether two letters are *semantically* identical or not (e.g., semantically identical: “Aa” or “AA” vs. semantically different: “Ab” or “AB”). In a meta-analysis, Neubauer (1995) reported an average correlation of  $-.33$  between mean reaction times and psychometric intelligence test scores. This suggests that brighter individuals display a higher speed of information processing than less intelligent individuals, probably due to a more efficient functioning of their central nervous systems (Jensen 1998). A central restriction of ECTs represents the rather low effect sizes of the observed correlations. In most cases, correlations do not exceed absolute values of  $.30$ ; a recent meta-analysis reports a mean correlation of  $-.24$  (Sheppard and Vernon 2008). Thus, mental speed usually accounts for scarcely more than 10 percent of the variance in intelligence tests.

The second potential basis of individual differences in cognitive abilities is *working memory*. Working memory (WM) can be regarded as a limited-capacity system responsible for temporary storage (or maintenance) and processing of information (Baddeley 2002; 2003). The inclusion of a processing component distinguishes WM from short-term memory (STM) which only supports temporary storage of information. As an example, in a prototypical STM task (forward span), two to nine words are presented sequentially, and the participants are required to recall the words afterwards in the same order. WM tests usually require the execution of a second, additional task. In the original reading span task, for instance, participants read aloud sentences while trying to remember the last word of each sentence for later recall (Daneman and Carpenter 1980). Individuals differ in the capacity of WM, and these differences have proven to be related to several higher-order cognitive functions ranging from rather domain-specific skills (like reading comprehension; Daneman and Carpenter 1980; vocabulary learning; Daneman and Green 1986; or numeracy; De Rammelaere et al. 1999) to (domain-general) intelligence. The actual size of correlation between WM capacity and intelligence as well as the appropriate statistical approach to determine their true relationship are matters of intensive debate (Ackerman et al. 2005; Beier and Ackerman 2005; Kane et al. 2005; Oberauer et al. 2005). The current estimates range between about  $.40$  and  $.80$ ; single previous studies reported even higher correlations (up to  $.96$ ) which led some authors to conclude that WM may be the psychological mechanism underlying (fluid) intelligence (Kyllonen and Christal 1990; Colom et al. 2004).

The distinction between storage and processing is also reflected in cognitive theories of WM. Probably the most prominent theory was put forward by Baddeley and colleagues already in the 1970s (Baddeley and



Hitch 1974). According to their tripartite model, WM consists of two “slave systems” which are coordinated and controlled by a third system, the “central executive.” The slave systems enable the temporary storage of information and are domain-specific: phonologically coded material (verbal and numerical material) is maintained in the phonological loop, visuo-spatial information in the visuo-spatial sketchpad. The central executive component was considered to be an attention-control mechanism which is responsible for focusing attention to (task-) relevant information, dividing attention if two tasks are performed, and switching attention between different processes and information (Baddeley 2002).

There is considerable evidence that the central executive component of WM is domain-independent and drives the relationship between WM capacity and intelligence (e.g., Engle et al. 1999; Kane et al. 2004; but see also Colom et al. 2005). More specifically, Conway and colleagues (2003) regarded the “active maintenance of goal-relevant information in the face of interference” (p. 549) as the critical cognitive basis that is shared between intelligence and WM tasks. Support for their view comes from findings that individuals with high and low WM capacity also differ in the performance of low-level attention-control tasks that place practically no memory demands on the participants. In the anti-saccade task, for example, participants have to make an eye movement (saccade) in the opposite direction of a visual cue (e.g., a flashing light in the periphery). Since the reflexive response would be to orient towards the cue, the attention control demand consists of suppressing this habitual response. Individuals with higher WM capacity were found to display faster and more correct saccades than individuals with lower WM capacity.

## 2. Status quo

At present, numerous psychometric “intelligence tests” are available. Virtually all of the currently available market tests do a good job with measuring individual differences in cognitive abilities in that they meet the main criteria required for a psychometric test: objectivity, reliability, and validity.

A test displays *objectivity* if the result is independent of the person who administers, analyzes, and interprets the participant’s performance. Objectivity is ensured by standardized instructions during administration as well as by clear-cut instructions for how test scores are determined and interpreted.

*Reliability* builds upon objectivity and reflects the measurement precision of a test. Reliability is never perfect (1.0) as the test performance is not only influenced by the true cognitive ability of the person but also by random factors such as momentary fluctuations of attention or mood, fatigue,

etc. Usually, intelligence tests display reliabilities around .90, indicating that 10 percent of the total variance in test performance is due to random factors (i.e., measurement error) and 90 percent reflects true variance in intelligence.

Finally, the *validity* of a test reflects to what extent the test measures the trait or ability that it should measure. The validity of intelligence tests is typically evaluated by relating the performance in the test under investigation to an external criterion, either to the performance in a well-established intelligence test or to criteria such as school grades. The great success of the concept of intelligence primarily originates in the high validity of intelligence test performance for a lot of performance indicators in diverse areas of life (e.g., Jensen 1998).

In line with the originally intended purpose of intelligence tests, the strongest associations are found with educational variables. Intelligence correlates with school grades at about .50 and with years of education at about .55 (Neisser et al. 1996). Intelligence can also be regarded as a good predictor of vocational success; in a meta-analysis Schmidt and Hunter (1998) reported an average validity of .51 for overall job performance. Another quality criterion of psychometric tests is the availability of norms so that the individual test performance can be compared with the performance of an age-matched reference sample. The norms in intelligence tests allow the determination of the IQ, reflecting the standardized position of an individual relative to a reference population with a mean of 100 and a standard deviation of 15.

Given their high reliability and validity, intelligence tests can be definitively regarded as the best choice for assessing cognitive abilities. Many of the available market tests not only provide an estimate of the general intelligence of an individual (the IQ) but also inform about his or her cognitive ability structure. The Berlin Intelligence Structure Test (BIS-T) (Jäger et al. 1997), for instance, assesses three content facets (verbal, numerical, spatial-figural) and four operational facets (processing capacity, creativity, memory, and speed) of cognitive abilities with general intelligence as the integral of all ability facets.

The administration of such an intelligence structure test, however, is very costly, predominantly in terms of time. The full version of the BIS-T takes over 2 hours. But even one-dimensional intelligence tests focusing on general intelligence, such as the Raven's Advanced Progressive Matrices (Raven 1958) require a test time of at least 20-30 minutes in their short version. Thus, if we want to disentangle the impact of cognitive abilities and socio-economic effects on the outcomes of human lives, there is a strong need for the development of shorter cognitive ability assessment procedures that can be applied in large-scale surveys.

Lang and colleagues (2007) recently proposed two ultra-short tests for the measurement of intellectual abilities in the German Socio-Economic

Panel (SOEP, *Sozio-oekonomisches Panel*). One test (the symbol-digit test, or SDT) requires the fast assignment of numbers to symbols following a pre-defined number-symbol pairing. In the other test (the Animal Naming Task, or ANT), participants have to produce as many animal names as possible within a 90 second time interval. The reliabilities of these tests were reported to be around .90 for the SDT and around .65 for the ANT respectively. Their validities for general intelligence however, were not investigated, but can be expected to be rather low. The SDT draws on mental speed, and the performance in similar task versions was found to be only weakly related to intelligence (Conway et al. 2002). Likewise, the ANT only samples knowledge in a certain domain which turned out to be correlated only between .33 and .39 with broader vocabulary knowledge (Lang et al. 2007).

### 3. Future developments

In consideration of the recent theoretical insights into the cognitive bases of intelligence and the consistently strong relationship between WM capacity scores and higher-order intellectual abilities, it seems very promising to further develop short tests that draw on WM or its sub-components. In contrast to intelligence problems, WM tasks typically require only simple cognitive operations whose sequence is highly restricted by the instructions. The difficulty of working-memory tasks arises from the additional load on some facets of the cognitive architecture (Süß et al. 2002). The reading-span task described above, for example, requires continuous updating of the content of WM (with every sentence one new word needs to be memorized) and the maintenance of the words in spite of interference (i.e., reading sentences aloud).

Overall, WM tests offer the following advantages:

- (1) They take a shorter time to administer than intelligence tests.
- (2) Most of the tasks involved can be implemented in computer-aided testing environments.
- (3) According to current research, they tap the central basis of cognitive abilities.
- (4) WM tasks are typically less influenced by prior knowledge than intelligence tests.
- (5) The limiting factor of WM capacity (central executive) seems to be domain-independent.

At this point, the development of WM tests lags far behind the development of intelligence tests. WM span tasks (such as the reading span task described above) are among the first WM measures to have been developed and are already well understood, a fact reflected in the existence of methodological reviews and user guides (Conway et al., 2005). The psychometric quality of other WM tasks (e.g., focusing on executive processes) is more difficult to evaluate due to the scarcity of studies with larger samples. In the following, an overview of WM tasks that could be employed in the large-scale survey studies is provided.

### 3.1 Traditional WM span tasks

Since the early reading-span task described above, several versions of WM span tasks have been developed. Three key tasks can be identified (Conway et al. 2005; Kane et al. 2004). In the (newer version of the) *reading span task*, the participant is presented with a meaningful or meaningless sentence and a to-be-remembered letter (e.g., “We were fifty lawns out at sea before we lost sight of land. ? X”). The participant’s task is to read the sentence, judge whether it makes sense or not, read, and remember the letter. The *operation span task* requires judging the correctness of an arithmetic equation and to remember an additionally presented word (e.g., “Is  $(6 \times 2) - 5 = 7$ ? class”). In the *counting span task* participants have to count the number of dark blue circles in displays with other distracting objects (dark blue squares and green circles) and to remember the counted number. All these tasks are designed to force storage of information in the face of processing.

Conway et al. (2005) emphasized three critical task features: first, rehearsal must be avoided by presenting the next stimulus immediately after completion of the preceding one. Second, the timing of the task needs to be adaptive. Both properties are met in current computer versions in which the to-be-remembered stimulus is displayed immediately after completion of the interfering task (e.g., judging the correctness of an equation). Third, the number of stimuli within one item needs to be sufficient. A range from two to five stimuli per item turned out to be adequate for most college students.

The administration of a WM span task with 12 items (with two to five stimuli each) including instruction and practice items takes about 10 minutes. Besides the verbal WM span measures described above, a number of figural-spatial versions have been devised (Kane et al. 2004). As an example, in the symmetry span task, participants have to judge whether a figure in an 8 x 8 matrix is symmetrical or not and to remember the position of a red square in a subsequently presented 4 x 4 matrix.

The reliabilities of WM span tasks are usually in the range between .70 and .90, suggesting good measurement precision for a single test. Their validity for intelligence test performance lies around .50 (Kane et al. 2004).

### *3.2 Transformation span tasks*

In this type of WM task, participants are not required to simultaneously store and process information but rather to perform some mental transformation on the stored information. A promising example is the alpha span task, originally developed by Craik (1986). Three to seven words are successively presented to the participant who is required to memorize them. After presenting the last word, the participant has to repeat the first letter of each word in alphabetical order, thus requiring an alphabetical reordering of the memorized words. Süß et al. (2002) presented one item with three words and two items with four, five, six, and seven words each, requiring an estimated test time of about 5 minutes including instruction.

The authors reported a reliability of .81 and a validity for general intelligence of .55. Other studies, however, report much lower validities for similar transformation tasks (e.g., the backward span task requiring the recall of the presented words in reverse order; Engle et al. 1999).

### *3.3 Dynamic WM tasks*

A separate class of WM tasks that are frequently used in neuroscience research require the continuous monitoring and updating of the maintained information. In the prominent n-back task, a list of stimuli (words, numbers, or figures) is successively presented, and the individual has to continuously report whether each stimulus matches the one that had appeared *n* items ago (n-back). In a 2-back task, for instance, participants have to continuously maintain the last 2 stimuli of the list which means that they have to update the content of their WM with every new stimulus and to drop out the least recent one. Even though the n-back task is considered the gold standard in neuroscience research, there is mixed empirical evidence on the question whether this task draws on the same cognitive resources as the well-established WM span tasks (Conway et al. 2005; Kane et al. 2007).

Kane et al. (2007) investigated the construct validity of the n-back task in a sample of 129 young adults and found that the performance in the operation span task and the n-back task was only weakly associated (correlations did not exceed .25). In addition, both tasks accounted for independent variance in general intelligence. These findings suggest that the n-back task does not measure the same WM processes as the operation span task.

### *3.4 Executive control tasks*

Executive processes related to attentional control are central in Baddeley's model of WM and are assumed to play a critical role in the relationship between WM capacity and intelligence. The development of tasks demanding these processes without strong reliance on storage, however, appears to be a great challenge. Süß et al. (2002) as well as Oberauer et al. (2003) have devised tasks requiring task set switching (i.e., the inhibition of an active action schema and the selection of another). In the numerical switching task by Süß et al. (2002), displays with varying number of digits are presented. The participant is required to alternate between reading the digits and counting them; the specific task to be performed is displayed on the top of the display. In the figural version, a round and an angular figure appears in each display, one left and one right. Participants have to indicate the side of either the angular or the round figure. Finally, in the verbal version, participants have to switch between two semantic categories in determining the presentation side of words. Similar to the transformation span tasks, these tasks can be administered within a few minutes.

Süß et al. (2002) report reliabilities between .78 (numerical) to .94 (verbal and figural) and validities between .33 (figural) and .58 (numerical) for general intelligence. Later research, however, has questioned the construct validity of these tasks as they are only weakly related to traditional WM span tasks (Oberauer et al. 2003; 2005) and reflect processing speed more strongly than reasoning abilities (Süß et al. 2002).

## **4. Conclusions and recommendations**

In the past decades, considerable advances have taken place in understanding the individual differences in cognitive abilities and in the development of psychometric tests for ability assessment. Present research regards WM, reflecting a limited-capacity system supporting temporary storage and processing of information, as the cognitive key system underlying intellectual abilities.

Measures of WM capacity have been found to display substantial correlations with several domain-specific intellectual abilities as well as with intelligence, representing the epitome of domain-general cognitive abilities. Thus, tests assessing WM capacity or executive functions appear to be a more promising method for the cognitive ability assessment in large-scale survey studies than tests focusing on mental speed or surface knowledge in a certain domain.

Several candidate tasks have been described above which can be administered in considerably shorter time than psychometric intelligence tests. In addition, their task characteristics allow the presentation in computer-aided testing environments. The Internet seems to offer the ideal infrastructure for the implementation of the cognitive ability screening. The coverage is very high, and it is meanwhile not longer only accessible from the personal computer (at home or at the office) but increasingly also from mobile devices such as netbooks, mobile phones, or personal digital assistants (PDAs). So it becomes ever more unproblematic to administer those tests in large-scale surveys.<sup>1</sup>

However, it should be noted that most of these WM tasks are still in the development phase, and that studies with larger samples, which would allow a more accurate evaluation of their reliability and (construct) validity, are very scarce. Thus, some initial steps would be very helpful. Although the future challenge is to improve the psychometric quality of these tests, they also need to be administered in large-scale surveys. In fact, the data of the large-scale surveys can further contribute to their improvement. The actual reliability of these tests could be accurately quantified and norms for age-matched reference samples, which are presently almost completely missing for WM tests, could be easily established. In addition, the data from large-scale studies can also inform about their validity for indicators of life success. Parallel to these criteria, their validity for intelligence needs to be further investigated.

---

1 The tests could be offered and advertized, for instance, in virtual social networks such as Facebook.

## References:

- Ackerman, P.L./Beier, M.E. and Boyle, M.O. (2005): Working memory and intelligence: The same or different constructs? *Psychological Bulletin* 131, 30-60.
- Baddeley, A.D. (2002): Is working memory still working? *European Psychologist* 7 (2), 85-97.
- Baddeley, A.D. (2003): Working memory: Looking back and looking forward. *Nature Reviews Neuroscience* 4 (10), 829-839.
- Baddeley, A.D. and Hitch, G.J. (1974): Working memory. In: Bower, G.A. (Ed.): *Recent advances in learning and motivation*, Vol. 8. New York.
- Beier, M.E. and Ackerman, P.L. (2005): Working memory and intelligence: Different constructs. Reply to Oberauer et al. (2005) and Kane et al. (2005). *Psychological Bulletin* 131 (1), 72-75.
- Binet, A. (1905): New methods for the diagnosis of the intellectual level of sub-normals. *L'Année Psychologique* 12, 191-244. [Translated in 1916 by Kite, E.S. in: *Development of intelligence in children*. Vineland, NJ: Publications of the Training School at Vineland].
- Brody, N. (1999): What is intelligence? *International Review of Psychiatry* 11 (1), 19-25.
- Carroll, J.B. (1993): *Human cognitive abilities: A survey of factor analytic studies*. Cambridge.
- Colom, R./Flores-Mendoza, C./Quiroga, M.A. and Privado, J. (2005): Working memory and general intelligence: The role of short-term storage. *Personality and Individual Differences* 39 (5), 1005-1014.
- Colom, R./Rebollo, I./Palacios, A./Juan-Espinosa, M. and Kyllonen, P.C. (2004): Working memory is (almost) perfectly predicted by g. *Intelligence* 32 (3), 277-296.
- Conway, A.R.A./Cowan, N./Bunting, M.F./Therriault, D.J. and Minkoff, S.R.B. (2002): A latent variable analysis of working memory capacity, short-term memory capacity, processing speed, and general fluid intelligence. *Intelligence* 30 (2), 163-183.
- Conway, A.R.A./Kane, M.J. and Engle, R.W. (2003): Working memory capacity and its relation to general intelligence. *Trends in Cognitive Sciences* 7 (12), 547-552.
- Conway, A.R.A./Kane, M.J./Bunting, M.F./Hambrick, D.Z./Wilhelm, O. and Engle, R.W. (2005): Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review* 12 (5), 769-786.
- Craik, F.I.M. (1986): A functional account of age differences in memory. In: Klix, F. and Hagendorf, H. (Eds.): *Human memory and cognitive capabilities*. Amsterdam.
- Daneman, M. and Carpenter, P.A. (1980): Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior* 19, 450-466.
- Daneman, M. and Green, I. (1986): Individual-differences in comprehending and producing words in context. *Journal of Memory and Language* 25, 1-18.
- De Rammelaere, S./Stuyven, E. and Vandierendonck, A. (1999): The contribution of working memory resources in the verification of simple mental arithmetic sums. *Psychological Research-Psychologische Forschung* 62 (1), 72-77.



- Deary, I.J. (1998): Differences in mental abilities. *British Medical Journal* 317, 1701-1703.
- Deary, I.J./Whiteman, M.C./Starr, J.M./Whalley, L.J. and Fox, H.C. (2004): The Impact of Childhood Intelligence on Later Life: Following up the Scottish Mental Surveys of 1932 and 1947. *Journal of Personality and Social Psychology* 86 (1), 130-147.
- Diewald, M. (2008): Zwillings- und Adoptivkinder-Stichproben für soziologische Analysen? Zum Ertrag verhaltensgenetischer Ansätze für sozialwissenschaftliche Fragestellungen und Erklärungen [Using twin and adoptive child samples for sociological analysis? On the value of behavioral genetics approaches for social scientific issues and interpretations]. *DIW Research Notes* 27.
- Engle, R.W./Tuholski, S.W./Laughlin, J.E. and Conway, A.R.A. (1999): Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology* 128 (3), 309-331.
- Gardner, H. (1983): *Frames of mind*. New York.
- Grigorenko, E.L. (2007): How Can Genomics Inform Education? *Mind, Brain, and Education* 1 (1), 20-27.
- Guo, G. (2008): Introduction to the special issue on society and genetics. *Sociological Methods & Research* 37 (2), 159-163.
- Gustafsson, J.-E. (1984): A unifying model for the structure of mental abilities. *Intelligence* 8, 179-203.
- Jäger, O.A./Süß, H.M. and Beauducel, A. (1997). *Berliner Intelligenzstruktur-Test [Berlin Intelligence Structure Test]*. Göttingen.
- Jensen, A.R. (1998): *The g factor: The science of mental ability*. Westport.
- Kane, M.J./Conway, A.R.A./Miura, T.K. and Colflesh, G.J.H. (2007): Working memory, attention control, and the N-back task: A question of construct validity. *Journal of Experimental Psychology-Learning Memory and Cognition* 33 (3), 615-622.
- Kane, M.J./Hambrick, D.Z. and Conway, A.R.A. (2005): Working memory capacity and fluid intelligence are strongly related constructs: Comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin* 131, 66-71.
- Kane, M.J./Hambrick, D.Z./Tuholski, S.W./Wilhelm, O./Payne, T.W. and Engle, R.W. (2004): The generality of working memory capacity: A latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology: General* 133, 189-217.
- Kyllonen, P.C. and Christal, R.E. (1990): Reasoning ability is (little more than) working memory capacity?! *Intelligence* 14, 389-433.
- Lang, F.R./Weiss, D./Stocker, A. and von Rosenblatt, B. (2007): Assessing Cognitive Capacities in Computer-Assisted Survey Research: Two Ultra-Short Tests of Intellectual Ability in the German Socio-Economic Panel (SOEP). *Schmollers Jahrbuch* 127 (1), 183-192.
- Neisser, U./Boodoo, G./Bouchard, T.J./Boykin, A.W./Brody, N./Ceci, S.J./Halpern, D.F./Loehlin, J.C./Perloff, R./Sternberg, R.J. and Urbina, S. (1996): Intelligence: Knowns and unknowns. *American Psychologist* 51, 77-101.
- Neubauer, A. (1995): *Intelligenz und Geschwindigkeit der Informationsverarbeitung*. Wien.

- Oberauer, K./Schulze, R./Wilhelm, O. and Süß, H.M. (2005): Working memory and intelligence – their correlation and their relation: Comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin* 131, 61-65.
- Oberauer, K./Süss, H.M./Wilhelm, O. and Wittman, W.W. (2003): The multiple faces of working memory: Storage, processing, supervision, and coordination. *Intelligence* 31 (2), 167-193.
- Posner, M.I. and Mitchell, R.F. (1967). Chronometric analysis of classification. *Psychological Review* 74 (5), 392-409.
- Raven, J.C. (1958): *Advanced Progressive Matrices*. London.
- Schmidt, F.L. and Hunter, J.E. (1998): The Validity and Utility of Selection Methods in Personnel Psychology: Practical and Theoretical Implications of 85 Years of Research Findings. *Psychological Bulletin* 124 (2), 262-274.
- Sheppard, L.D. and Vernon, P.A. (2008): Intelligence and speed of information-processing: A review of 50 years of research. *Personality and Individual Differences* 44 (3), 535-551.
- Spearman, C. (1904): "General intelligence", objectively determined and measured. *American Journal of Psychology* 15, 201-293.
- Spinath, F.M. (2008): Improvements and Future Challenges in the Field of Genetically Sensitive Sample Designs. RatSWD Working Paper No. 45.
- Stern, W. (1912): *Die psychologischen Methoden der Intelligenzprüfung*. Leipzig.
- Süß, H.-M./Oberauer, K./Wittmann, W.W./Wilhelm, O. and Schulze, R. (2002): Working memory capacity explains reasoning ability – and a little bit more. *Intelligence* 30, 261-288.
- Wechsler, D. (1944): *Measurement of adult intelligence*. Baltimore.

## 2.2 Measuring Cognitive Competencies

*Ulrich Trautwein*

Contact:

Ulrich Trautwein  
University of Tuebingen  
Institute of Education  
Europastraße 6  
72072 Tuebingen  
Germany  
e-mail: [ulrich.trautwein\[at\]uni-tuebingen.de](mailto:ulrich.trautwein[at]uni-tuebingen.de)

## **Abstract**

The systematic assessment of key cognitive competencies is of great scientific and societal interest, as is the availability of high-quality data on cognitive competencies. In order to make well informed decisions, politicians and educational authorities need reliable data about the effectiveness of formal and non-formal educational environments. Similarly, researchers need strong data to test complex theoretical models about how individual biographies are shaped by the interplay between individual and institutional affordances and constraints.

There are countless datasets that offer some form of information on competencies, such as the respondents' years at school and their school grades. Such data are relatively easy to collect. When it comes to making informed political and educational decisions, however, there are increasing calls for a more systematic use of standardized competency tests. Yet the production, storage, and use of standardized test data on competencies in specific domains is expensive, complex, and time-consuming.

This advisory report argues that there is a paucity of adequate data on cognitive competencies in important domains, and especially a lack of longitudinal data from standardized competency tests. Moreover, in the case of many important questions there are no good alternatives to high-quality standardized tests of cognitive competencies. Finally, it outlines some challenges in the construction and application of standardized competency tests and makes several recommendations.

Keywords: cognitive competencies, assessment, intelligence, school grades

## **1. The need for systematic assessments of cognitive competencies**

Competencies are abilities that allow an individual to master the complex demands of particular contexts. There are many components of competent performance, including knowledge, cognitive and practical skills, attitudes, emotions, values, and motivations (for a detailed definition of competencies, see Rychen and Salganik 2001; Weinert 2001). The scope of this report is restricted to cognitive competencies that are taught and learned in formal and non-formal learning environments. These cognitive competencies include, for instance, subject-specific knowledge, reading and mathematical literacy, computer literacy, and job-related knowledge.

The systematic and rigorous assessment of key cognitive competencies is of high scientific and societal relevance, as is the availability of high-quality data on cognitive competencies. It is now widely accepted that in modern knowledge societies the economic prosperity of individuals, communities, and countries is associated with the cognitive competencies an individual has

acquired. In order to make well informed decisions, politicians and educational authorities need high-quality data about the effectiveness of formal and non-formal educational environments. Similarly, researchers need strong data to test complex theoretical models about how individual biographies are shaped by the interplay between individual and institutional opportunities and constraints. Some questions that require high-quality data on cognitive competencies include: has mathematics and reading literacy generally increased or decreased among high school students in recent years and decades (see Becker et al. 2006)? Are *Abitur* (the school-leaving certificate in the academic track) standards and related competence levels comparable across the German federal states, or *Länder* (Trautwein et al. 2007)? Do female and male students and students from different family and ethnic backgrounds have the same access to high-quality education? Which domain-specific competencies are important for success in different domains at university and in the workplace (e.g., Nagy 2006)?

There is general agreement about the importance of assessing and documenting the competence levels achieved by learners in formal and non-formal learning environments; innumerable datasets offer some form of information on competencies. For instance, official statistics report the number of students who leave school with specific school-leaving certificates; school authorities document the distribution of school grades assigned within different grade levels and school types in each school year; and scientific studies ask students about their academic standing relative to their peers.

Such data on school-leaving certificates, grades attained in various learning environments, and self-reports of achievement are relatively easy to collect. They can be used to inform many questions and add to the body of knowledge about educational systems. When it comes to making informed political and educational decisions, however, there are a growing number of calls for a more systematic use of standardized competency tests. In the wake of international benchmarking studies such as PISA (OECD Programme for International Student Assessment) (Baumert et al. 2001), there has been increased scientific interest in Germany in the conceptualization, psychometric modeling, operationalization, and description of cognitive competencies. The production, storage, and use of standardized test data on competencies in specific domains is expensive, complex, and time-consuming, however. The question therefore arises of whether standardized competency tests might be replaced by cheaper and more readily available alternatives.

This report argues that there is a paucity of adequate data on cognitive competencies in important domains. There is, in particular, a lack of longitudinal data from standardized competency tests, and for many important questions there are no good alternatives to high-quality standardized tests of cognitive competencies. This chapter is structured as follows: the following

section provides a short description of standardized cognitive competency tests. Section 3 highlights the differences between domain-specific cognitive competencies and intelligence tests. Three frequently used but qualitatively different approaches to measuring cognitive competencies (grades or certificates; self-reports of competence; self-concepts) are compared and contrasted in section 4, and their advantages and disadvantages discussed. Challenges in the construction and application of standardized competency tests are subsequently outlined. Finally, several recommendations are made. For the sake of brevity and drawing from available data, the chapter draws primarily on data collected in schools and universities.

## 2. Standardized tests of cognitive competencies

Standardized tests of cognitive competencies use student responses to certain stimuli (or “items”) to infer competence levels. Carefully constructed standardized assessments such as those used in the PISA study are based on a conceptual model of what is being assessed, and their construction and evaluation is informed by psychometric models and state-of-the-art statistical analyses. Psychometrically constructed standardized tests have to fulfill a number of criteria. Most importantly, they must be objective (i.e., the resulting test scores must be independent of the person who administers and scores the test), reliable (i.e., the test must be internally consistent and give consistent results over time), and valid (i.e., the test must actually measure what it sets out to measure). When standardized tests with high validity are used, the competence levels of all test takers can be compared directly, independent of where they live or their learning environment. Some well-known standardized tests of cognitive competencies include the TOEFL test assessing the English-language skills of non-native speakers and the PISA tests assessing verbal, mathematical, and scientific literacy.

Tests can be distinguished along several dimensions. *Curriculum-oriented tests* are based on material defined in the learners’ curriculum. For instance, a curriculum-oriented mathematics test would implement tasks covered in the mathematics curriculum. In contrast, tests such as those implemented in PISA, which are based on the *literacy concept*, probe for competencies considered essential for full participation in society. Ideally, tests of cognitive competencies allow comparison across test takers (“*norm-referenced tests*”) and inform on the individual test taker’s absolute level of competence (“*criterion-referenced tests*” or tests with “*competence levels*”).

Standardized tests such as PISA have helped to close the knowledge gap surrounding the cognitive competencies of various student groups (e.g., boys vs. girls; students with different immigration status). For instance, findings

have shown that immigrant groups are differentially successful in different school systems. When carefully constructed, standardized tests have a high degree of fairness because all students receive a similar “treatment”. Moreover, it is possible to discern items that may place some subgroups at a disadvantage and to eliminate these items from the test.

A specific advantage of carefully constructed standardized competency tests is that they allow the development of competence to be tracked over time. Forms of “anchoring” allow test scores to be compared longitudinally, provided that the conceptual model is good and the quality of measurement is high.

To date, longitudinal data on the development of cognitive competencies over time are in short supply in Germany. Although some datasets contain such information, they tend to be relatively small, restricted to some areas of Germany, and/or the competency tests used are of limited quality (Blossfeld 2008). The National Educational Panel Study (NEPS; Blossfeld 2008) commencing in 2009 will help to overcome this deficit by examining students’ mathematics, verbal, scientific, and ICT literacy as well as their literacy in English as a foreign language. Furthermore, some domain-specific tests will be administered to subgroups (e.g., business students).

### **3. Cognitive competencies vs. intelligence**

The construction of psychometrically sound tests of domain-specific cognitive competencies is complex and expensive. Some critics have questioned whether these efforts are strictly necessary or whether cheaper alternatives are available. One proposed alternative is to use measures of general, decontextualized cognitive dispositions, such as intelligence (Rindermann 2006). Rindermann claimed that the competency tests used in large-scale assessment studies such as TIMSS (Trends in International Mathematics and Science Study) and PISA measure a single cognitive ability that is practically identical to general intelligence. Given the relatively high intercorrelations observed between mathematics literacy, reading literacy, and cognitive ability, it might therefore be argued that it would be easier and cheaper to use intelligence tests instead of tests of domain-specific competencies in large-scale assessments.

This line of argumentation has major limitations, however (Baumert et al. 2007). First, there are clear conceptual differences between domain-specific cognitive competencies and general, decontextualized cognitive dispositions such as intelligence (e.g., in processes of knowledge acquisition and information processing and in dependence on the quality of educational environments). Second, although there is a statistically significant correlation



between intelligence and scores on domain-specific competency tests, the results of construct validation studies provide strong empirical support for the multidimensionality (i.e., empirical separability) of cognitive measures applied in large-scale educational assessments (see Baumert et al. 2007). Third, evaluations of the educational effectiveness of a specific school, state, or country differ across domains, as shown, for instance, by a recent study (Trautwein et al. 2007) comparing educational outcomes at the end of the academic track in two German states (Baden-Württemberg and Hamburg). Although the Baden-Württemberg students clearly outperformed the Hamburg students in mathematics, with an effect size of Cohen's  $d = .98$ , the respective differences in English achievement ( $d = .16$ ) and reasoning ( $d = .07$ ) were negligible. Fourth, intelligence and domain-specific competencies differentially predict academic outcomes such as success at university (Nagy 2006).

Taken together, domain-specific cognitive competencies are theoretically and empirically separable from general, decontextualized cognitive dispositions such as intelligence, which are less amenable to educational interventions (see the expertise by Stern, in this publication). Tests of intelligence cannot replace psychometric tests of cognitive competencies in assessments of educational effectiveness.

## **4. Other measures of cognitive competencies**

### *4.1 Grades and (school-leaving) certificates*

Many datasets contain information on teacher-assigned school grades and/or (school-leaving) certificates. For example, official statistics in Germany document in detail a broad range of certificates acquired in formal education (e.g., school-leaving certificates; university diplomas; completed apprenticeships). Similarly, many datasets contain information on teacher-assigned grades or teacher evaluations of student progress (e.g., school grades; university grades). Without question, grades and certificates affect individuals' academic biographies and long-term success on the job market, and thus represent important information that should be documented. However, to what extent can these easily available data replace information obtained using complex and expensive standardized achievement tests? Three aspects are critical here: reference group effects, the association between background variables and teacher-assigned grades, and the reliability of self-reports.

#### 4.1.1 Reference group effects: restricted comparability

Prior research has clearly documented that achievement scores collected via standardized achievement tests correlate only moderately with teacher-assigned school grades (Baumert et al. 2003; Ingenkamp 1971). Although teacher-assigned grades typically give a rather accurate (but not perfect) estimate of the position of each student within a class, teachers' differential grading standards mean that grades do not typically provide a valid basis for gauging achievement across classes or schools. The majority of teachers in Germany and in many other education systems do not use an absolute criterion for achievement when assigning grades (as is the case in standardized achievement tests). Rather, they tend to grade on a norm-referenced basis (Ingenkamp 1971), with the best student in the class receiving a very good grade and the weakest student a bad grade or a "fail." As a consequence, "grading-on-a-curve" effects can be observed in most schools in Germany. The size of the correlation between school grades and standardized competency tests typically ranges from about  $r = .30$  to  $r = .60$  (e.g., Baumert et al. 2003; Trautwein et al. 2007). When individual achievement is controlled, higher class-average achievement is associated with lower grades (Trautwein et al. 2006). Clearly, it is important to distinguish theoretically and empirically between these two indicators of achievement. Furthermore, teacher-assigned grades cannot easily be used to measure learning gains over time.

Given that school grades are not on a common metric across teachers and schools, it is hardly surprising that students who acquire the same school-leaving certificate in different schools or states do not necessarily exhibit the same level of cognitive competencies. Moreover, although qualifications such as the *Hauptschulabschluss* (lower-school certificate) and *Abitur* (certificate giving access to higher education) are awarded at various school types in Germany, little is known about the comparability of these certificates across school types.

#### 4.1.2 Effects of sex, family background, and immigration status

Teacher-assigned grades have been shown to be influenced not only by cognitive competencies, but also by various student characteristics. Importantly, it is well documented that teacher evaluations of students' cognitive competencies are associated with the students' family backgrounds (e.g., Baumert et al. 2001) and influenced by teachers' gender stereotypes. Teachers' evaluations of students from immigrant families are also likely to be affected by stereotypes, but these effects may be compensated by grading leniency. More empirical studies are needed in this context.

#### 4.1.3 Validity of self-report data

Information on school grades and certificates can be collected via either self-reports or school records. Are student self-reports of their grades reliable indicators of their actual grades or should school records be consulted? Several recent studies have reported high associations between self-reported and teacher-reported grades. For instance, Dickhäuser and Plenter (2005) reported a correlation of  $r = .88$  for the last mathematics grade. It must be noted, however, that participants in these studies did not have anything to gain from reporting higher school grades than they actually attained. In a different context, the association between self-reported and teacher-reported grades may well be lower.

#### 4.2 *Self-assessments of cognitive competencies: restricted validity and group differences*

A quick, easy, and direct approach to assessing cognitive competencies is to ask individuals for an “objective” evaluation of their own competencies. For instance, students might be asked to report their competencies in logical reasoning or grammar (e.g., Kruger and Dunning 1999) and these self-ratings are then correlated with data from a standardized test or an expert rating. As shown in a meta-analysis by Mabe and West (1982), the resulting associations are typically moderate in magnitude and vary from study to study. Mabe and West were able to identify some characteristics of studies that moderate the association between self-ratings and other indicators of competencies. Higher associations are found, for instance, if respondents expect their self-reports to be compared with objective evaluations and if some guarantee of anonymity is given in the study instructions. Even under such favorable conditions, however, the associations between self-reported competencies and external information on these competencies were far from perfect. Looking at various domains, moreover, Kruger and Dunning (1999) showed that people with low abilities in these domains were particularly likely to overestimate their abilities. Kruger and Dunning attributed these misjudgments partly to the lower metacognitive competencies of these respondents. Taken together, the validity of self-evaluations of cognitive competencies is restricted.

#### 4.3 *Domain-specific self-concepts*

Domain-specific self-concepts are another frequently used construct in many empirical studies. Domain-specific academic self-concepts reflect a person's self-evaluation regarding a specific academic domain or ability (see Traut-

wein et al. 2006). These self-concepts are usually collected via self-report measures. Typical self-concept items are “I am quite good at mathematics” (mathematics self-concept) and “I have a poor vocabulary” (verbal self-concept). Although self-concepts share some similarities with self-evaluations of competencies, there is one crucial difference. Self-concept instruments ask specifically for a person’s subjective self-evaluation, not for an “objective” self-evaluation. It is therefore not surprising that these instruments elicit external frame of reference effects (e.g., respondents compare their accomplishments with those of their friends or schoolmates rather than using an “average” comparison group) as well as internal frame of reference effects (e.g., respondents compare their competencies in mathematics with their competencies in English), yielding a complex pattern of associations with other assessments of competencies. Domain-specific self-concepts have proven to be predictive with regard to the competency development. However, they are no substitute for standardized tests of cognitive competencies.

#### *4.4 The need for multiple indicators: a research example*

Which indicator of cognitive competence is the best predictor of a successful transition from school to university or the labor market? Modern educational systems work on the assumption that competence levels predict future success in higher education and the workforce. However, it has also been argued (e.g., Solga 2005) that employers rely heavily on the type of school-leaving certificate as a “signal” when hiring apprentices or employees. These certificates are more easily accessible than, for instance, test scores, and may thus have more influence in determining applicants’ professional success than their actual level of competence. There is indeed some reason to believe that – given their easy availability to employers – school-leaving certificates and school grades have more pronounced effects on success in the application process, whereas competencies predict success during vocational training and occupational careers. As plausible as this reasoning may seem, however, there is a need for empirical studies that empirically tease apart the confounding effects of certificates and competence levels cross-sectionally and longitudinally. Moreover, there is a need to distinguish among facets of cognitive abilities. Some studies from the United States seem to indicate that general ability (intelligence) plays a more important role in training success than do specific competencies (e.g., Ree and Earles 1991). Convincing empirical support for such a pattern of results is still lacking in the German context, however, primarily because of the lack of datasets including information on competencies measured by standardized tests as well as teacher-assigned grades and certificates.

## 5. Standardized competency tests: Challenges

The empirical assessment of competencies is more difficult than it may appear at first glance. Theoretically and empirically sound competence models are required as a basis for the development of measurement procedures. The systematic integration of theoretical frameworks, psychometric models, and measurement approaches often requires interdisciplinary cooperation, which introduces another level of complexity. The challenges facing longitudinal competence measurement outlined below are among those currently being addressed.

As yet, there is disagreement over which *domains of cognitive competencies* can be meaningfully measured by standardized tests and how differentiated the measurement should be. These questions are, for instance, very relevant to job-related cognitive competencies. Similarly, with regard to *criterion-referenced tests and competence levels*, there is also some disagreement over which competence levels can be considered sufficient, which levels can the majority of learners realistically achieve, and who should be responsible for establishing these standards in different domains.

Another challenge pertains to *possible positive and negative effects of competence testing*. What are the effects of systematic competence assessment in learning environments? For instance, do teachers make changes to the learning content covered or to their methods of teaching in response to the introduction of competency tests, and are the overall effects positive or negative? Moreover, to date, in standardized educational assessments such as PISA, *unmotivated test taking* might have been the exception rather than the rule in Germany (Baumert and Demmrich 2001). However, it is not clear if this may change in the future if standardized competency tests are administered more frequently.

Finally, *longitudinal measurement* is one of the most difficult and crucial challenges in the context of competence testing. Challenges include choosing an appropriate linking procedure, possible retest effects, the danger of ceiling and floor effects, and the question of whether the construct being measured remains the “same” over time (e.g., are multiplication tables in elementary school and complex numbers at upper secondary level part of the “same” mathematics?).

## 6. Recommendations

- (1) Broader use of standardized tests of cognitive competencies can help to evaluate the effectiveness of educational institutions. Wherever feasible, standardized competency tests – in addition to or instead of measures such as teacher-assigned grades and self-assessed competence – should be used.
- (2) Some effort should be put into investigating domains of cognitive competencies for which competency tests can be easily constructed (based on either the curriculum or the literacy concept) and domains for which standardized competency tests are not feasible.
- (3) There is a need for more high-quality tests that are available for researchers for use in their own projects (e.g., intervention studies). In this sense, there should be broader access not only to data but also to measurement instruments.
- (4) When reporting “competence” data, researchers should always specify whether standardized tests or alternatives were used. It is especially important to critically address possible reference group effects and whether specific groups (e.g., gender or immigrant groups) might be at a disadvantage.
- (5) Concerted efforts should be made to strengthen expertise in constructing and interpreting standardized competency tests in the scientific and non-scientific communities. There has been considerable progress in recent years (e.g., the German Research Foundation’s (DFG, *Deutsche Forschungsgemeinschaft*) priority program on “Competence models for recording individual learning outcomes and for reviewing educational progress”), but more expertise is needed across a broader population of researchers.
- (6) Ways must be found of linking competence data collected in empirical studies (e.g., school achievement studies) with other datasets (e.g., data available from national agencies).

## References:

- Baumert, J./Brunner, M./Lüdtke, O. and Trautwein, U. (2007): Was messen internationale Schulleistungsstudien? – Resultate kumulativer Wissenserwerbsprozesse. Eine Antwort auf Heiner Rindermann [What do international student achievements measure? – Results of cumulative learning. A reply to Heiner Rindermann.] *Psychologische Rundschau* 58, 118-129.
- Baumert, J. and Demmrich, A. (2001): Test motivation in the assessment of student skills: The effects of incentives on motivation and performance. *European Journal of Psychology of Education* 16, 441-462.
- Baumert, J./Klieme, E./Neubrand, M./Schiefele, U./Schneider, W./Stanat, P./Tillmann, K.-J. and Weiß, M. (Eds.) (2001): PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich [PISA 2000: Student literacy in international comparison]. Opladen.
- Baumert, J./Trautwein, U. and Artelt, C. (2003): Schulumwelten: Institutionelle Bedingungen des Lehrens und Lernens [School environments: Institutional conditions of learning and instruction]. In: Baumert, J./Artelt, C./Klieme, E./Neubrand, M./Prenzel, M./Schiefele, U./Schneider, W./Tillmann, K.J. and Weiß, M. (Eds.): PISA 2000: Ein differenzierter Blick auf die Länder der Bundesrepublik Deutschland. Opladen.
- Becker, M./Trautwein, U./Lüdtke, O./Cortina, K.S. and Baumert, J. (2006): Bildungsexpansion und kognitive Mobilisierung [Educational expansion and cognitive mobilization]. In: Hadjar, A. and Becker, R. (Eds.): Die Bildungsexpansion. Erwartete und unerwartete Folgen. Wiesbaden.
- Blossfeld, H.-P. (2008): Education as a lifelong process. A proposal for a National Educational Panel Study (NEPS) in Germany. Bamberg.
- Dickhäuser, O. and Plenter, I. (2005): "Letztes Halbjahr stand ich zwei": Zur Akkuratheit selbst berichteter Noten [On the accuracy of self-reported school marks]. *Zeitschrift für Pädagogische Psychologie* 19, 219-224.
- Ingenkamp, K. (1971): Die Fragwürdigkeit der Zensurengebung [The dubiousness of school grades]. Weinheim.
- Kruger, J. and Dunning, D. (1999): Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology* 77, 1121-1134.
- Mabe, P.A. and West, S.G. (1982): Validity of self-evaluation of ability: A review and meta-analysis. *Journal of Applied Psychology* 67, 280-296.
- Nagy, G. (2006): Berufliche Interessen, kognitive und fachgebundene Kompetenzen: Ihre Bedeutung für die Studienfachwahl und die Bewährung im Studium [Vocational interests, cognitive and scholastic abilities: Their role in choice of major and success at university]. Unpublished doctoral dissertation, Freie Universität Berlin.
- Ree, M.J. and Earles, J.A. (1991): Predicting training success: Not much more than g. *Personnel Psychology* 44, 321-332.
- Rindermann, H. (2006): Was messen internationale Schulleistungsstudien? Schulleistungen, Schülerfähigkeiten, kognitive Fähigkeiten, Wissen oder allgemeine Intelligenz? [What do international student assessment studies measure? School

- performance, student abilities, cognitive abilities, knowledge, or general intelligence?]. *Psychologische Rundschau* 57, 69-86.
- Rychen, D.S. and Salganik, L.H. (2001): Defining and selecting key competencies. Göttingen.
- Solga, H. (2005): Ohne Abschluss in die Bildungsgesellschaft: Die Erwerbschancen gering qualifizierter Personen aus soziologischer und ökonomischer Perspektive [Entering the education society without qualifications. The employment opportunities of lower qualified individuals from the sociological and economic perspectives]. Opladen.
- Trautwein, U./Köller, O./Lehmann, R. and Lüdtke, O. (Eds.) (2007): Schulleistungen von Abiturienten: Regionale, schulformbezogene und soziale Disparitäten [School achievement of students at Gymnasium schools: Regional, track-specific, and social disparities]. Münster.
- Trautwein, U./Lüdtke, O./Marsh, H.W./Köller, O. and Baumert, J. (2006): Tracking, grading, and student motivation: Using group composition and status to predict self-concept and interest in ninth grade mathematics. *Journal of Educational Psychology* 98, 788-806.
- Weinert, F.E. (2001): Vergleichende Leistungsmessung in Schulen – eine umstrittene Selbstverständlichkeit [Comparative measurement of achievement in schools]. In: F. E. Weinert (Ed.): *Leistungsmessungen in Schulen*. Weinheim.



## 2.3 Measuring Vocational Competencies

*Martin Baethge and Lena Arends*

Contact:

Martin Baethge  
Sociological Research Institute (SOFI)  
University of Goettingen  
Friedlaenderweg 31  
37085 Goettingen  
Germany  
e-mail: martin.baethge[at]sofi-uni-goettingen.de

## **Abstract**

International large scale assessment programmes mostly concentrate on measuring and comparing general competences of students in compulsory education or adults in everyday life; concepts for measuring young adults' competences in specific vocational domains are lacking. The paper addresses the need for competence measurement in vocational education and training and illustrates future demands for the implementation of an international large-scale assessment of vocational education and training (VET-LSA). It discusses possible measurement approaches for vocational competences and outlines why methods for measuring internal conditions (dispositions) by tests instead of external measurement by observation are the favourite option. The measurement of cross-country differences in the level and distribution of vocational competences could break new grounds and provide standardised and internationally comparable indicators for benchmarking of VET and a basis for adequate classification of learning outcomes in classification systems, such as the EQF.

Keywords: international large scale assessment, vocational education and training, VET, competence measurement, VET-LSA, large-scale assessment in VET

## **1. The political dimension of measurement in Vocational Education and Training (VET)**

The increasing internationalization of labor markets and vocational education and training (VET), as well as recent initiatives to improve transparency in VET such as the proposals for a European Qualification Framework (EQF) and a European Credit System for Vocational Education and Training (ECVET), have shifted the debate on transparency and quality standards in VET to the policy level.

The European Commission has put forward the ambitious economic and social goal of becoming "the most competitive and dynamic knowledge-based economy in the world" (European Commission 2004). In the field of VET this aim is being pursued through the Copenhagen Process. The European Commission's proposal for a EQF offers opportunities to increase mobility and enhance permeability between educational sectors. The recent progress report by the European Commission outlining advancement towards the Lisbon objectives emphasized that "internationally comparable large scale assessments programmes often concentrate on general competences (e.g., reading, information processing, numeracy and problem solving) whereas many employers argue that [...] there is an increasing need to conduct surveys which focus as well on the assessment of vocational skills and competences" (European Commission 2008: 61).

There are no indicators for international benchmarking of VET (see Higher Education below). The International Standards Classification of Education (ISCED, UNESCO 2006) does not provide a tool for comparing the performance of VET systems in different countries because the levels for classifying different VET programs are not transparent enough<sup>1</sup> (Mueller and Klein 2008).

An internationally comparable, objective, valid, and reliable assessment of cross-country differences in the level and distribution of vocational competences could provide a scientific basis for the debates on the EQF and the ECVET. The results would allow for a better classification of national educational programs within existing classification systems (e.g., NQF/EQF, ISCED) on an empirical basis.

## **2. Current state of competence measurement in VET**

### *2.1 Large-scale assessments in education*

There is a general consensus that indicators for measuring quality are key instruments for improving the quality of education and training necessary for the good governance of education systems and structures (European Commission – DG EAC 2004). Internationally comparative studies (e.g., PISA,<sup>2</sup> IALS<sup>3</sup>) provide standardized and internationally comparable indicators that give insight into the factors influencing the development of competences of students in compulsory education and adults in everyday life. These indicators provide policy-makers with a tool on which to base future policy choices.

The measurement of competences has become an instrument for benchmarking the performance of educational systems. During the past fifteen years, considerable experience has been gained in the development and implementation of large-scale assessments in education. International tests have been developed for measuring basic skills of students in compulsory education and generic skills of adults in everyday life; however, tests for measuring competences in specific vocational areas are still missing.

There are three major surveys for the measurement of basic skills of students in compulsory education: the TIMSS<sup>4</sup> survey, which focuses on the mathematics and science aptitudes of students in the fourth and the eighth

---

1 In particular ISCED 3B/3C, 4 and 5B.

2 The Programme for International Student Assessment (PISA) started in the mid 1990s.

3 The International Adult Literacy Survey (IALS) was administered in 1994, 1996, and 1998.

4 The Trends in International Mathematics and Science Study (TIMSS) was implemented first in 1995.

grade in a number of countries; the PIRLS<sup>5</sup> survey, in which reading literacy among fourth graders is measured; and the PISA study, which assesses the literacy, numeracy, science, and problem-solving performance of fifteen-year-olds. PISA has become a major policy tool for measuring student competencies in an internationally comparative perspective, for establishing benchmarks of educational improvement, and for representing the strengths and weaknesses of national educational systems.

For measuring the skills of adults, two surveys have been implemented: the IALS survey was the first attempt to implement a large-scale adult skills assessment based on tests on an international basis (OECD and Statistics Canada 2005). The Adult Literacy and Lifeskills (ALL) Survey builds on the IALS survey by extending the range of competency domains and improving the quality of assessment methods (Murray and Clermont 2005). The ALL Survey<sup>6</sup> is the most comprehensive program for measuring adult skills (16–65 years) in the following domains:

- prose literacy,
- document literacy,
- numeracy, and
- problem solving.

The ALL framework is based on an all-encompassing skill typology involving a comprehensive groups of skills relevant for everybody in a variety of life contexts (e.g., community, home, work). Skills for work are based on employability skill models derived from general skill categories that are applicable across a variety of working situations.

Building on the ALL survey, the OECD has launched an international study of financial literacy of adults in twelve OECD countries. The assessment is focused on financial literacy in the context of everyday life among adults as consumers, workers, and citizens (OECD 2005; 2008).

Following up on the IALS and ALL surveys, the OECD is also currently developing a strategy for assessing the literacy skills of adults, including familiarity with information and communication technologies, and the ability to manage information, construct new knowledge, and communicate with others (Schleicher 2008). In the first cycle, measurement in PIAAC<sup>7</sup> will be implemented in three areas:

---

5 The Progress in International Reading Literacy Study (PIRLS) was implemented in 2001.

6 The assessment was conducted in 2003 in people's homes in seven countries: Bermuda, Canada, Italy, Norway, Switzerland, United States, Nuevo León.

7 Programme for the International Assessment of Adult Competencies. PIAAC is currently being implemented with a view to undertake a first assessment in 2011. The main survey design is based on a minimum main study sample size per country of 5,000 adults, aged 16–64.

- *Direct assessment of cognitive skills.* Integrated measurement of broad literacy competence encompassing the range of performance from mastery of the basic building blocks of literacy to the capacity of effectively managing complex information processing tasks embedded in an Information Technology (IT) setting. Up to four areas of competences will be assessed, including “Problem-solving in a technology-rich environment,” “Literacy,” “Numeracy,” and “Literacy component skills.”
- *Indirect assessment of skills used in the workplace.* The JRA<sup>8</sup> module seeks to assess the level and use of a number of generic skills such as communication, presentation, and teamwork skills in the workplace. With the JRA module, individuals will be asked a set of questions relating to their job and the requirements of that job in terms of the intensity and frequency of the use of certain skills.<sup>9</sup>
- *Background questionnaire.* The aim here is to collect information relating to the antecedents and outcomes of adult literacy competences, including information on literacy practices and familiarity with and usage of information technology.
- None of the illustrated surveys for cross-country differences measuring the level and distribution of competences of students and adults is focused on the assessment of domain-specific competences in specific vocational areas.

### *Approaches for measurement in VET*

The term “competence” is one of the most internally diverse concepts in the field of education and educational policy. Efforts to define the term have remained unsuccessful because a variety of topics are addressed under the heading of “competence” (e.g., Descy and Tessaring 2005; Arends 2006). In some contexts “competences” refer to the internal conditions of individuals that allow them to master tasks in different situations successfully. In others, “competences” refer to the tasks and situational requirements themselves. To compare the performance of different VET systems and programs, we differentiate between these two approaches for measuring competences: the measurement of the *internal conditions* that allow the individual to master (occupational) tasks successfully, and the measurement of the individual’s

---

<sup>8</sup> Jobs Requirement Approach (JRA).

<sup>9</sup> Given that the approach is untested in an international survey, a pilot of the JRA module is currently taking place to test the validity and reliability of this approach in a cross-country and cross-cultural context.

*external performance* in different (occupational) situations. Following PISA, we are in favor of approaches that measure internal conditions in different occupational fields. This approach perceives individuals as carriers of competences who have developed the ability and willingness to successfully apply their knowledge, abilities, and experiences to authentic (occupational) situations. To measure internal conditions, however, internationally valid models of the structure and development of competences in specific domains (e.g., literacy, numeracy in PISA) at different levels of complexity are needed.<sup>10</sup>

The development and implementation of internationally valid tests for measuring internal conditions is considered a time-consuming, expensive, and methodologically challenging task (e.g., item development in PISA). To overcome this, many skills surveys have tried to identify less time-consuming and cheaper approaches to measuring competence on the basis of external performance. Typically, external measurement is implemented through the direct observation of individuals during work performance (e.g., work samples) or self-assessment of work activities (e.g., JRA approach). For example, the UK Skills Survey (Ashton et al. 1999; Felstead et al. 2002) and the O\*NET database (US Occupational Information Network) are based on the Job Requirement Approach (JRA), a self-assessment instrument for individuals to rate their competences at the workplace.

Toolsema (2003) adopted the Generalized Work Activities (GWA) concept from O\*NET<sup>11</sup> for identifying and comparing competences. In O\*NET, GWA are “aggregations of similar job activities/behaviors that underlie the accomplishment of major work functions” (Jeanneret et al. 1999: 106). The central assumption is that work behavior is not necessarily linked to specific tasks and techniques and, therefore, can be located at a higher level of aggregation. Activities are considered competence indicators at a higher level of abstraction, indicating the purpose of competences. Toolsema derived six competence categories – social, participative, cognitive, physical-technical, learning, and employability – and linked them with GWAs. The instrument was applied for measuring cognitive, affective, and meta-cognitive aspects of performance at different levels of abstraction in the field of Higher Education, not in VET.

A different external approach was used in the VQTS model.<sup>12</sup> Here, the aim was to provide a structured description of work-related competences and their acquisition. By using a competence matrix, which displays the competences of a specific occupational field at different levels, the model aims to provide an instrument for VET providers to transfer and recognize competen-

---

10 In VET, models for measuring competences in different vocational areas are currently being developed (e.g., Winther and Achtenhagen 2008; Nickolaus et al. 2008).

11 Occupational Information Network from the US.

12 Leonardo da Vinci project "VQTS – Vocational Qualification Transfer System".

ces acquired within the official VET systems in foreign countries (Luomi-Messerer and Markowitsch 2006; Markowitsch et al. 2006). The model can be used for descriptions of qualification requirements in different occupational fields, but not for measuring individuals' abilities to successfully master occupational tasks in different occupational situations.

Approaches for external measurement of vocational competences are not recommendable for measuring competences in specific vocation areas:

- (1) Self-assessment instruments, such as JRA or GWA, produce a picture of individual activities within a particular workplace, not of the competences used to successfully master occupational tasks in different situations.<sup>13</sup> Individual competences cannot be inferred from descriptions of performance in different working environments because work activities are linked to concepts of specific work organizations, which differ between firms and occupational sectors as well as between national settings.
- (2) The greatest disadvantage of self-assessment as a method of obtaining data is the greater chance of measurement error, resulting from a more or less "intentional" manipulation of answers by respondents and unintentional discrepancies between the real and reported values, resulting in low reliability. To measure and compare young adults' performance across countries in a valid and reliable way, objective measures of internal conditions are highly recommendable.
- (3) Performance-based measurement of specific, workplace-related tasks must be based on an agreement of the performance levels in different occupations in terms of occupational tasks, and must be relevant in specific vocational areas across countries.
- (4) The overarching goals of VET are not sufficiently incorporated into external approaches: although aspects of work-related qualifications are included, those pertaining to individual development and societal participation are not.
- (5) Some external approaches are very time-consuming, which is problematic in large-scale assessments in terms of reaching valid, reliable, and objective results at a low cost (e.g., skills tests in the World Skills competition amount to twenty hours total testing time).
- (6) The most striking criticism that has been made of external measurement is the fact that data of students' performance cannot be linked to the organization and contents of VET and that it is not possible to assess the development of individual learning during VET.

---

13 GWAs could be used for triangulations in a large-scale assessment of VET.



### 3. Future demands for competence measurement in VET

To use competence measurement as an instrument for international benchmarking of VET systems and programs, we need an international large-scale assessment based on concepts for measuring internal dispositions using Item Response Theory models (Rost 2006). During its council presidency, the German government (Federal Ministry of Education and Research) adapted the discussion on quality in VET to launch an initiative for an “International Large-Scale Assessment of Vocational Education and Training” (VET-LSA).<sup>14</sup> The main benefit of a VET-LSA is to expand the knowledge for steering VET processes at different policy levels. The proposed VET-LSA will increase valid and reliable steering knowledge:

- to determine the relationship between individual (biographic) characteristics, training forms, and skill building;
- to improve transparency with regard to the performance of European VET programs;
- to link VET outcomes and institutional orders of VET systems;
- to determine the correlation between the competences certified in final examinations and competences actually measured;
- to assess the comparative strength and weaknesses of different training forms in different countries to learn from each other;
- to classify different vocational training qualifications in international classification schemes (ISCED, EQF) in order to support the comparability of certification processes at the European level;

and others.

The development and implementation of this type of international large-scale assessment of VET is much more complex than it is for compulsory education. Whereas international student assessment programs (e.g., PISA) are based on well-grounded research traditions and internationally validated concepts (e.g., world curriculum for mathematics, standards in education), VET cannot draw upon any similarly comparable concepts. In a VET-LSA, the focus is on the measurement of competences in specific occupational domains rather than on all occupations. Contrary to international assessment programs of students and adults (e.g., PISA, PIAAC), VET-LSA does not

---

14 On the policy level, the following representatives from European countries and institutions are members in the international Steering Group: Sweden, Denmark, Norway, Slovenia, Finland, Switzerland, Germany, Austria, Spain, CEDEFOP, ETF, EC.

claim to be an overall representative survey. The aim is to rather include some of the most important industrial–technical, commercial–administrative, and care occupations in the sample (Baethge et al. 2008; Baethge et al. 2007; Baethge et al. 2006).

VET systems in different countries or even in sectors within the same country are unlikely to be drawn into convergence through the Copenhagen Process (2002); there is a variety of outcomes rather than conformity. To measure the same construct, common elements in VET have to be identified at the outset. A VET-LSA feasibility study, which is currently being developed in cooperation with experts from all participating countries,<sup>15</sup> will provide a clear picture of those national programs that are comparable and might be included in an international comparison (Baethge et al. 2008). As a first step in making comparisons, international classification systems (ISCED, ISCO, O\*NET) are used as a frame of reference for identifying a set of core occupational tasks and qualification requirements within all participating countries, not as a tool for competence measurement.

One of the central requirements for an international comparison of VET with a focus on competence measurement is a common understanding of the appropriate objectives for VET. This common understanding must be mutually developed, taking into consideration different scientific and policy perspectives. The definition of objectives for VET can be based on either a relatively limited or a broader approach. A broader approach encompasses individual competences in specific occupational domains and competences that individuals need to participate effectively as members of a flexible, adaptable, and competitive workforce and as lifelong learners. In a limited approach, competences are rather focused on the requirements of the workplace. In accordance with the ongoing scientific discussion, three objectives for educational systems at the systemic level were determined with an international group of VET researchers (Baethge et al. 2007). These objectives function as reference points for the definition of competences in vocational education and training and for the development of measurement tools in a VET-LSA:

- First, the development of individual vocational adjustment from an individual user perspective that takes into account the critical aspect of autonomy in working situations (individual perspective).
- Second, the ability to deal with today's labor market requirements and develop one's career (human resources perspective).
- Third, the ability to participate in organizational processes of work and work-related interactions (social perspective).

---

15 Sweden, Denmark, Norway, Finland, Slovenia, Switzerland, Austria, and Germany. France and Spain are interested in participating in the VET-LSA.

To determine to what extent VET systems are able to achieve these objectives, they have been operationalized within the concept of the VET-LSA in cooperation with experts from different countries. The concept of the VET-LSA comprises two major areas of measurement:

- *Vocational and occupational domain-specific competences* refer to young adults' abilities to successfully apply their knowledge and experience to authentic occupational situations in specific vocational fields (Car Mechanics, Electricians, Business & Administration, Social & Health Care).
- *Cross-occupational competences* are related to successful performance in the labor market. They refer to the notion of key skills or "core competencies," which comprise knowledge, skills, and abilities, for example, about the structures of organizations and labor markets, or acting autonomously in work environments.<sup>16</sup>

To illuminate interdependencies between domain-specific vocational and basic competencies, a third area is included in the concept for VET-LSA:

- *Basic competences*, such as reading, writing, and mathematics, which are being tested in international programs for student or adult assessments.

In addition to the measurement of young adults' competences in VET, institutional and individual contextual factors impacting the development of competences during VET are included in the concept for the VET-LSA:

***Institutional conditions*** on a macro and meso level refer to

- coordination and steering (actors in VET and their responsibilities, e.g., state or social partners),
- standards and norms (e.g., curricula and exams),
- resources (e.g., financing, professionalism of teaching staff), and
- institutional cooperation of educational service providers.

Whereas institutional conditions of VET should be measured in terms of average conditions in occupational fields (not at the institutional level) on the basis of expert interviews, organizational conditions of VET should be measured in terms of young adults' perceptions derived from questionnaires.

---

16 The measurement approach for cross-occupational competences in VET-LSA differs from the approach for generic workplace skills in PIAAC. Whereas in PIAAC assessment is based on self-assessment of work activities with the JRA approach, cross-occupational competences in VET-LSA will be assessed on the basis of approaches for internal conditions.

***Individual conditions*** of educational development and background consist of:

- socio-economic status of the family (including migration),
- social and cultural capital of young adults and their families,
- educational and career development,
- information behavior and learning activities (including non-formal learning activities), and
- educational or career aspirations.

For the quality of educational processes, the following aspects are considered important:

- a focus on problem solving and task complexity within occupational contents,
- variability (multiple working tasks and tools),
- possibilities for independent task solving,
- support,
- learning climate,
- team integration, and
- quality control.

An international large-scale assessment of vocational education and training will provide a number of indicators to compare and evaluate VET systems and institutional arrangements in VET according to quality, including young adults' competence levels in different competence domains, and information about occupations and their relation to individual and institutional factors, among others.

## References:

- Arends, L. (2006): Vocational competencies from a life-span perspective: Theoretical considerations and practical implications for an international large-scale assessment of vocational education and training. Norderstedt.
- Ashton, D./Davies, B./Felstead, A. and Green, F. (1999): *Work Skills In Britain*. Oxford.
- Baethge M./ Arends, L. and Winther, E. (2008): First international workshop Feasibility Study VET-LSA, Discussion paper, workshop 3.-4. July, 2008 in Bonn. SOFI Göttingen.
- Baethge, M./Achtenhagen, F. and Arends, L. (2008): How to Compare the Performance of VET Systems in Skill Formation. In: Mayer, K.-U. and Solga, H. (Eds.): *Skill Formation – Interdisciplinary and Cross-National Perspectives*. New York.
- Baethge M./Achtenhagen, F./Arends, L./Winther, E./Nickolaus, R./Gschwendtner, T./Onna, M. van/Béguin, A./Sibberns, H. and Watermann, R. (2007): Concept for an International Large Scale Assessment of Vocational Education and Training (VET-LSA). Working paper for the international workshop, 29.-30. October, 2007 in Bonn. SOFI Göttingen.
- Baethge, M./Achtenhagen, F./Arends, L./Babic, E./Baethge-Kinsky, V. and Weber, S. (2006): *PISA-VET. A Feasibility-Study*. Stuttgart.
- Copenhagen Declaration (2002): Declaration of the European Ministers of Vocational Education and Training, and the European Commission, convened in Copenhagen on 29 and 30 November 2002, on enhanced European cooperation in vocational education and training. <http://ec.europa.eu/education/pdf/doc125en.pdf>.
- Descy, P. and Tessaring, M. (2005): *The value of learning: evaluation and impact of education and training: third report on vocational training research in Europe: synthesis report*. Luxembourg: Office for official publications of the European Communities (Cedefop reference series; 61).
- European Commission (2004): *Maastricht communiqué on the future priorities of enhanced European co-operation in vocational education and training: a review of the Copenhagen declaration of 30 November 2002*. Brussels.
- European Union (2008): *The Bordeaux Communiqué on enhanced European cooperation in vocational education and training*. [http://www.eua.be/fileadmin/user\\_upload/files/Newsletter\\_new/Bordeaux\\_Communique\\_EN.pdf](http://www.eua.be/fileadmin/user_upload/files/Newsletter_new/Bordeaux_Communique_EN.pdf). [Last visited: 03/10/2010].
- Felstead, A./Gallie, D. and Green, F. (2002): *Work Skills In Britain 1986-2001*. Nottingham.
- Helsinki Communiqué (2006): *The Helsinki Communiqué on Enhanced European Cooperation in Vocational Education and Training*. Communiqué of the European Ministers of Vocational Education and Training<sup>1</sup>, the European Social partners and the European Commission, convened in Helsinki on 5 December 2006 to review the priorities and strategies of the Copenhagen Process. <http://ec.europa.eu/education/policies/2010/doc/helsinkicomen.pdf>. [Last visited: 03/10/2010].
- Jeanneret, P.R./Borman, W.C./Kubisiak, U.C. and Hanson, M.A. (1999): Generalized Work Activities. In: Peterson, N.G./Mumford, M.D./Borman, W.C./Jeanneret,

- P.R. and Fleishman, E.A. (Eds.): An occupational information system for the 21st century: The development of O\*NET. Washington.
- Luomi-Messerer, K. and Markowitsch, J. (Eds.) (2006): VQTS model. A proposal for a structured description of work-related competences and their acquisition. Vienna: 3s research laboratory.
- Markowitsch, J./Becker, M. and Spöttl, G. (2006): Zur Problematik eines European Credit Transfer System in Vocational Education and Training (ECVET). In: Grollmann, P./Rauner, F. and Spöttl, G.; (Eds.): Europäisierung Beruflicher Bildung – eine Gestaltungsaufgabe. Hamburg.
- Mueller, W. and Klein, M. (2008): Von Schmollers-Vorlesung 2008: Schein oder Sein: Bildungsdisparitäten in der europäischen Statistik. Eine Illustration am Beispiel Deutschlands. Schmollers Jahrbuch 128 (4), 511-543.
- Murray, T.S. and Clermont, Y. (2005): The origins and objectives of the ALL study. In Murray, T.S./Clermont, Y. and Brinkley, M.: Measuring adult literacy and life skills. New frameworks for assessment. Ottawa.
- Nickolaus, R./Gschwendtner, T. and Geißel, B. (2008): Modellierung und Entwicklung beruflicher Fachkompetenz in der gewerblich-technischen Erstausbildung. Zeitschrift für Berufs- und Wirtschaftspädagogik 104 (1), 48-73.
- OECD (2005): Improving financial literacy: Analysis of issues and policies. Paris.
- OECD (2008): Improving financial education and awareness on insurance and private pensions. Paris.
- OECD and Statistics Canada (2005): Learning a Living. First results of the adult literacy and life skills survey. Paris.
- Rost, J. (2006): Using Item-Response-Theorie for measuring vocational competences in an international large-scale assessment in VET – Expert report. In: Baethge, M./Achtenhagen, F./Arends, L./Babic, E./Baethge-Kinsky, V. and Weber, S. (Eds.): PISA-VET. A Feasibility-Study. Stuttgart.
- Schleicher, A. (2008): PIAAC: a new strategy for assessing adult competencies. International Review of Education.
- Toolsema, B. (2003): Werken met competenties. Naar een instrument voor de identificatie van competenties. Enschede.
- UNESCO (2006): ISCED 1997 International Standard Classification of Education. Re-edition. <http://www.uis.unesco.org/TEMPLATE/pdf/isced/ISCEDA.pdf>.
- Winther, E. and Achtenhagen, F. (2008): Kompetenzstrukturmodell für die kaufmännische Bildung. Adaptierbare Forschungslinien und theoretische Ausgestaltung [Competence model in the field of business and administration: state of the art and directions for further research]. Zeitschrift für Berufs- und Wirtschaftspädagogik 104 (4), 511-538.

## 2.4 Measuring Social Competencies

*Ingrid Schoon*

Contact:

Ingrid Schoon  
Institute of Education  
University of London  
Department of Quantitative Social Science  
20 Bedford Way  
London WC1H 0AL  
UK  
e-mail: I.Schoon[at]ioe.ac.uk



## **Abstract**

What are social competencies? How can they be measured? Do they remain stable over time? This report examines the difficulties encountered conceptualizing and measuring social competencies at different developmental stages and in changing social contexts. Existing measures and available data sources are reviewed and recommendations are made for future developments in data provision, data usage, and data access.

Keywords: social competence, social skills, social relationships and interaction

## **1. Measuring social competencies**

Social competencies have been identified by the European Commission as one of the benchmark indicators of prosperity and well-being that must be targeted for improvement in Member States (EU 2005). Social competencies can be broadly defined as the capabilities that enable individuals “to live together in the world” (Arendt 1958), comprising aspects of interpersonal, intercultural, social, and civic competencies. Beyond such a broad definition, however, general social competencies are difficult to define because the skills and behaviors needed for living together in the world and for achieving social tasks and outcomes vary with age and with the demands of particular situations. The notion of social competence is of interest to social scientists across disciplines, since it relates to adaptive functioning in a variety of contexts and across the lifespan. Social competencies reflect adjustment within the family, school, work, in society at large, and to old age. Therefore, more context-specific definitions of the construct are required, in addition to a greater focus on particular facets of social competence, such as empathy, self control, trust, respect for other people, and civic engagement. In recent years, the study of social competencies has received increased attention from policy-makers and social scientists across disciplines, partly due to increased concerns about the lack or erosion of social competencies in modern society (see for example Putnam 2000).

### *1.1 Conceptual issues*

A major concern for empirical research is that social competencies are generally not well defined or measured. Social competencies comprise interactions between individual characteristics, social demands, and situative characteristics. They have to be understood as relative, since very different kinds of social competencies are required and valued in different contexts

(Argyle et al. 1985). Behaviors that are functional in one context might be dysfunctional in another, implying that the assessment of social competencies involves culturally based value judgments. These values, however, are subject to change. For example, as a consequence of the massive economic transformations that have taken place there, new behaviors and qualities – such as assertiveness and autonomy – are now required in China to be successful, whereas characteristics that used to be beneficial for adjustment – such as obedience to authority – are now perceived as problematic (Chen and French 2008).

Social competencies are conceptualized differently across disciplines, and even within disciplines there is no consensus about their exact definition. Within psychology, for example, social competencies are defined variously as personality traits (Sarason 1981) that can manifest themselves in different capabilities such as empathy, tolerance, and conscientiousness; as the ability to cooperate; as a dynamic construct involving successful adjustment to and interaction in given social conditions (Argyle 1994; Tajfel 1981); as people's beliefs about their own efficacy (Bandura 1997), and as social (Gardner 1999) or even emotional intelligence (Goleman 1995). In a pedagogical context it refers to lifelong, intercultural, and social learning. In economics, social competencies are sometimes used to refer to "soft skills" comprising abilities such as flexibility, working in a team, and motivating colleagues and clients. Economic terms such as "social capital" are used in sociology and the social sciences in general to describe resources arising from social relationships (Putnam 2000; Halpern 2005). Given this variety of definitions it is necessary to establish a unifying working definition that acknowledges differences in focus, and which specifies particular domains of manifestation as well as specific components and skills. To avoid confusion researchers must be clear about their theoretical orientation and must identify the context and focus of their assessments.

## *1.2 Research questions*

*Development of competencies.* Social competencies change over the life course and depend on the development of capabilities such as social awareness, social skills, and self-confidence. For example, young children learn to play games with others, such as peek-a-boo or let's-pretend games, but in the process also learn important forms of self-control, including patience, sharing, and temper management, as well as learning to empathize with others. Later on they have to develop more integrated forms of self-regulation, with an emphasis on "fitting in" and achievement, as well as increased coordination of social skills and understanding of social scripts as they unfold (Saarni 2000; Waters and Sroufe 1983). Certain behaviors may be appropriate at particular ages but not at others. We still know relatively

little about the developmental antecedents, or about the outcomes of social competencies in areas such as health, well-being, socio-economic attainment, and social integration. Questions that need to be addressed include concerns with how social competencies are expressed at different periods during the life course, but also whether there are stages in life when it is too early to expect a sense of social responsibility or empathy. Are there particularly sensitive periods of heightened awareness? What is the potential for developing social competencies over the life course? To answer questions about the development and growth of social competencies, and to assess continuity and change in development over time, it is vital to be able to draw on longitudinal data that follows individuals from an early age onwards. Furthermore, there must be agreement on the key indicators of social competencies at different life stages based on a thorough theoretical understanding of human development in context.

*Biological aspects.* There may be links between social competencies and other enduring personality characteristics as well as genetic factors that shape social interactions. However, there is still little understanding of the association of social competencies and genetic factors or physiological measures of neural efficacy (Flashman et al. 1998; Grigorenko 2000; Bechara et al. 2000). Nor do we yet know whether there are some basic physical and psychological needs that have to be fulfilled before social competencies can be developed.

*Social change.* In recent years concerns have been raised about the erosion of social competencies as a consequence of socio-historical change and increasing globalization. Some have argued that there has been an increasing instrumentalization and individualization of social relationships (Putnam 2000), while others have emphasized the emergence of new values and lifestyles with greater tolerance for ethnic, cultural, and sexual diversity, more issue-oriented forms of participation, greater emphasis on self-expression, and a search for the meaning and purpose of life (Inglehart 1997). Until recently it has not been possible to analyze the linkages between macro-social change and individual level attitudes due to the lack of reliable time-series data measuring certain concepts repeatedly across many different societies, or large-scale longitudinal studies following the development of social competencies within individuals over time and across different birth cohorts. Today, however, a number of large-scale longitudinal studies following individual lives over time as well as international panel studies are accessible, such as the European Values Study and the World Values Survey (WVS), both of which have been used to test assumptions about changing social values and competencies (Arts and Halman 2004; Inglehart 1997).

*Context dependency.* Social competencies are essentially relational, describing how individuals behave within the context of interpersonal and group relationships. Characteristics of both the relationships in which indi-

viduals engage and their context offer opportunities to acquire and express social competencies. How are competencies influenced by interactions with family members, peers, at school or at work, or in one's neighborhood? What are the factors and processes that foster and promote social competencies? To answer these questions it is vital to assess information about contextual as well as individual characteristics. Questions about the transgenerational transmission of social competencies and values, a process not yet fully understood, require the assessment of social competencies across generations as well as a consideration of socialization practices and the availability of social support.

Another area of inquiry concerns general versus context-specific manifestations of social competence. To what extent can social competencies be generalized across groups and communities? How do opportunities, norms, and expectations for social connectedness and participation influence the development of social competencies over the life course? Key context-related indicators that must be considered include measures of social status (comprising socio-economic as well as family status, education, and income), gender, culture, and ethnicity, formal and informal settings, as well as age. Questions to be addressed by researchers include, for example, whether the gender or cultural differences that are often noted in the expression and/or manifestation of social competencies are a product of measurement, norms and socialization influences, or something else.

### *1.3 Measurement*

There is no established consensus about how to measure social competence. The assessment of social competencies can comprise a variety of methods, ranging from self-ratings or self-reports of behavior, values, and motivations; direct behavioral observations (in natural situations or under experimental conditions); behavior rating scales (to be completed by parents, teacher, employer, subordinates, or self); the use of vignettes; interviewing; make-believe tasks and role-play; hypothetical scenarios; the interpretation of video clips; social network analysis and sociometric approaches; to computer simulations.

A widely used instrument to assess personality characteristics such as agreeableness, conscientiousness, or extraversion is the "Big Five" inventory and its abbreviated forms (Costa and McCrae 1992; Gosling et al. 2003; McCrae and Costa 2004). Other widely used self-reported measures are the Rosenberg Self-Esteem Scale (Rosenberg 1979), measures of self-efficacy (Bandura 1997; Schwarzer 1993) or locus of control (Rotter 1990), or the Interpersonal Reactivity Index (Davis 1983) that measures both cognitive and affective aspects of dispositional empathy. Useful scales for the assessment of social adjustment in children and adolescents are the social competence

inventory (Rydell et al. 1997), the self-control rating scale (Kendall and Wilcox 1979), the child behavior checklist (Achenbach and Howell 1993), and the Strengths and Difficulties Questionnaire (SDQ), which contains subscales measuring peer problems and prosocial behavior (Goodman 2001). These questionnaires are by no means a complete list of available instruments. They are intended as examples of the many ways it is possible to conceptualize and operationalize social competencies. Generally it is best to select measures geared toward the context being addressed. There are also widely used single-item measures that are often included in large-scale surveys, which tap into conceptions of generalized trust (most people can be trusted), reciprocity, social networks and support, or social participation.

Concerns have been raised about the consistency and reliability of self-assessment as well as biases in reporting (Hagerty et al. 2007). Single-item measures, although attractive, are only suitable for the assessment of constructs that are simple and unambiguous. They are otherwise limited in several respects: they provide only one chance to capture a complex concept, they are likely to miss differences at the individual level, and they might be “contaminated” by the context in which they are collected. Psychometric scales comprising multiple items to measure a specific dimension – such as social intelligence, social responsibility, assertiveness, or empathy – are more reliable yet often take a longer time to complete and without abbreviation are not suitable for large-scale surveys. The same applies to attempts to measure social competencies on the basis of assessments in experimental settings, make-believe scenarios, or interpretations of video clips, which usually take more time to collect. A compromise might be to use or to develop brief multi-item scales for specific competencies. Another major concern is the lack of clarity or agreement on what indicators are relevant to establish construct validity. Definitions either focus on internal processes or external outcomes, although both aspects are important. Ideally the measurement of social competencies should involve different assessment modes – combining self-reports, rating scales completed by others, and observational data to obtain reliable and valid measures. Instead of direct assessments, multiple measures could be used as indicators of latent constructs, which would also facilitate comparative approaches of assessment and research.

## 2. Status quo: Databases and access

Free web-based access to national and international studies is available through a variety of social science data archives across Europe and the US. The UK based *Economic and Social Data Service (ESDS)* provides support for secondary use and facilitates access to an extensive range of key quantitative and qualitative economic and social data. The ESDS Qualidata archive provides access to qualitative data, such as the study “Inventing Adulthoods” that explores social relationships and interactions among young people living in the UK,<sup>1</sup> or the study, “Quality of Life in Older Age,”<sup>2</sup> providing information on social networks and support.

ESDS also assists users locating and acquiring international survey data,<sup>3</sup> as well as longitudinal data.<sup>4</sup> Data collections include, for example, the 1958 National Child Development Study (NCDS), the 1970 British Cohort Study (BCS70), the British Household Panel Survey (BHPS), the English Longitudinal Study of Ageing (ELSA), the Families and Children Study (FACS), the Longitudinal Study of Young People in England (LSYPE), and the Millennium Cohort Study (MCS). These studies contain a wide range of data on social competencies, comprising assessments in early childhood, adolescence, and adulthood. The MCS, for example, a study of over 18,000 children born between 2000-2002 includes measures of early social competence, using the Strength and Difficulties Questionnaire as well as a make-believe task (Sally-Anne task). NCDS and BCS70 contain measures of early behavioral adjustment, using the Rutter A Scale.<sup>5</sup> NCDS, a fifty-year-old study, contains measures of the NEO “Big Five” personality inventory. Most of the studies include assessments of social attitudes in adulthood, such as attitudes towards equality and fairness and information about social networks and civic activities – although mostly as single-item statements.

The *Council of European Social Science Data Archives (CESSDA)* is an umbrella organization for social science data archives across Europe. The CESSDA Portal<sup>6</sup> is a gateway to many kinds of research data and metadata, including, for example, international panel studies that have adopted a collaborative approach with other countries to provide comparative data. Studies accessible via this portal include the International Social Survey Programme (ISSP), the European Social Survey (ESS), the European Values Study, the World Values Surveys (WVS), and the International Social Justice Project (ISJP). All of these surveys contain items assessing generalized social trust

---

1 <http://www.esds.ac.uk/findingData/snDescription.asp?sn=5777>

2 <http://www.esds.ac.uk/findingData/snDescription.asp?sn=5237>

3 <http://www.esds.ac.uk/international/>

4 <http://www.esds.ac.uk/longitudinal/>

5 <http://www.data-archive.ac.uk/doc/5805/mrdoc/pdf/RutterBehaviourQuestions.pdf>

6 <http://damad.essex.ac.uk/portal/cessda.html>

(using a question such as “most people can be trusted”), frequency of contact with friends and relatives, strengths of social networks, participation in social and civic activities, social attitudes, attitudes towards gender equality, and social justice.

*The Inter-University Consortium for Political and Social Research (ICPSR)*, based at the University of Michigan<sup>7</sup> is an organization of member institutions working together to acquire and preserve social science data, to provide open and equitable access to these data, and to promote effective data use. ICPSR is the world’s largest archive of digital social science data. It provides, for example, access to the following longitudinal datasets that contain data on competencies, attitudes, values, and behaviors: the Panel Study of Income Dynamics (PSID), the National Longitudinal Survey of Youth 1979 and 1997 (NLSY79; NLSY97), data on the children of the National Longitudinal Survey of Youth (NLSYC), the National Longitudinal Study of Adolescent Health (Add Health), and Monitoring the Future. The NLSY studies, for example, contain information about self-esteem (Rosenberg scale), self-efficacy (Pearlin scale) of both mother and children, and information about behavior adjustment (Achenbach Youth Report), delinquency, social relationships, and social networks. The Child Development Supplement (CDS) of the PSID and the Add Health Study also contain information about self-esteem and self-efficacy as well as information on social support and social attitudes. The PSID CDS provides time-use diary data accounting for the social context of daily social activities. The Add Health Study contains information on dyadic relationships and social networks, enabling a close analysis of relationship symmetry, the strengths of friendship ties, and social integration.

### 3. Future developments

#### 3.1 Data provision

Given the stock of available data resources and the multiple perspectives on how to operationalize social competences, future challenges for data provision include (a) the integration and consolidation of existing data resources and measures of social competencies, (b) the cataloging and documentation of topic-specific resources, (c) the promotion of data re-use, (d) the addition of data sources to the archives that have not yet been made available, and (e) efforts toward the harmonization of future data collection.

*Integration and consolidation.* So far, very few attempts have been made to take stock of and evaluate existing resources. Future efforts should attempt

---

7 <http://www.icpsr.umich.edu/>

to provide an overview of existing measures and approaches. Similarities and differences in approach, as well as unifying conceptual issues, must be identified to enable the development of integrative research.

*Cataloging and documentation.* Currently documentation exists for individual studies (most of which are multipurpose), yet topic-specific documentation of measures and methodologies across different studies is lacking. Combining topic-specific evidence from different studies, countries, populations, and age groups will facilitate comparative research and contribute to a more integrated conceptualization of research. In particular, information about types of assessment, age groups and populations under investigation, psychometric properties of assessment (i.e., reliability and validity), inter-linked context variables, relevant publications, and the strengths and weaknesses of the particular approach used, are needed.

*Promotion of data re-use.* To date, not all relevant studies have been made available for public access via data depositories. This includes large-scale multipurpose longitudinal data, as well as more focused specialized investigations. To gain a better understanding of the different approaches and contexts of assessment, it is necessary to overcome “proprietary” models of publicly funded social science research and shift toward a more open and collaborative paradigm in order to obtain as complete an overview as possible, drawing on existing evidence. Of course, studies should be vetted and evaluated for criteria of research excellence before they are added to the depositories.

*Data harmonization.* Future data collection should build on the existing evidence base and strive to promote coordinated collaborative efforts following best practices, ideally involving several countries to provide comparative data.

### 3.2 Data usage

Data usage in the future is likely to involve interdisciplinary teams and international research networks sharing and consolidating existing knowledge, working towards a coordinated, comparative approach, and preparing strategies for collecting new evidence. To facilitate such developments it is necessary to improve the infrastructure of international data provision – including data documentation across different studies and disciplines and possibly the creation of innovative examples that model how to use data from different studies.



### *3.3 Data access*

Access to data should be expanded via remote access to sites and coordinated data archives. Consequently, resources must be invested in protecting the confidentiality of data, and consideration must be given to different levels of access depending on security clearance. Given the attractiveness of personal data for various interest groups and financial or market organizations, safeguarding the data while allowing access to bona fide users is a vital issue.

### *3.4 European and international challenges*

A key issue in the collaborative use of data is the international comparability of data and the provision of internationally harmonized datasets. Such an endeavor has to build on collaborative agreements between contributors and joint research projects. Every effort should be made to preserve existing data and to enable its re-use, even with a different purpose or research question in mind. Language barriers have to be acknowledged and overcome, for example, through coordinated efforts in data collection and documentation. Another area of importance concerns culturally specific norms and resulting expectations about what constitutes social competencies. For such issues it is necessary to identify a common denominator or to develop culturally sensitive or culturally neutral measures.

Furthermore, existing data sources should be integrated, creating multipurpose studies. This might involve the linkage of panel and cohort studies to administrative data and expanding the scope of studies to assess predictors and outcomes of development across domains, such as education and health. Innovative tools for data collection and analysis have to be developed, making use of newly available technologies. For example, data collection can be conducted via mobile phone or the internet, even using advanced methods of assessment such as computer simulation or time-use diaries. Further consideration should be given to the development of new analytic approaches, enabling the analysis of mechanisms and processes across and within domains, contexts, cohorts, countries, and over time. Moving beyond population statistics, there is the potential to adopt new methodologies that enable the identification of patterns and a comparison of functioning both between and within subgroups of the population.

### *3.5 Recommendations*

The measurement of social competencies involves the study of a complex phenomenon that occurs over time and in specific contexts. The following recommendations would advance our understanding and assessment of social competencies:

- Improved conceptual clarity and focus are needed to define what is being measured in the research. This should be achieved through efforts to develop an interdisciplinary, culturally sensitive, and relevant working definition of social competencies, integrating both general and specific components and skills.
- Appropriate methods are needed to map development over time and across domains and contexts. This implies the need for age-, domain-, and context-appropriate measures, enabling the assessment of growth and development over time. Methods should also be developed that are suitable for the examination of continuity and change in the acquisition and expression of competencies in different contexts.
- Since the effectiveness of social behavior can only be determined within the context of a particular social environment, it is necessary to include both individual and contextual characteristics in the assessment.
- The acquisition of social competencies is a developmental process, yet we have too little knowledge about how individuals learn and acquire social competencies in different contexts and settings, and how competencies develop and diversify over time. It is thus vitally important to increase the availability of longitudinal data, beginning at an early age to acquire information on different manifestations of competencies. Longitudinal data would also help clarify the factors and processes that facilitate the acquisition and expression of social competencies and promote adaptive interpersonal and personal environmental interactions at different life stages.
- There is a need for a better understanding of intergenerational transmission of social competencies as well as their biological foundations.
- To consolidate the research evidence, effort must be made to continuously update and advance the integration of existing data resources and to promote their re-use. Collaborative agreements to submit data to a publicly accessible data depository for the purpose of secondary analysis would pave the way for future shared research and training. Working toward the cataloging and topic-specific documentation of resources would provide the necessary infrastructure.

- To improve the possibility of collaborative and comparative research there should be integrated and harmonized approaches to data collection that draw on newly available technologies.
- Confidentiality of data has to be safeguarded, and specific modes of access to data depositories must be considered.

## References:

- Achenbach, Th.M. and Howell, C.T. (1993): Are American children's problems getting worse? *Journal of the American Academy of Child and Adolescent Psychiatry* 32 (6), 1145-1154.
- Arendt, H. (1958): *The Human Condition*. Chicago.
- Argyle, M. (1994): *The psychology of interpersonal behaviour*, 5<sup>th</sup> edition. London.
- Argyle, M./Henderson, M. and Furnham, A. (1985): The rules of social relationships. *British Journal of Social Psychology* 24, 125-139.
- Arts, W. and Halman, L. (Eds.) (2004): *European values at the turn of the Millennium*. Leiden.
- Bandura, A. (1997): *Self-efficacy: the exercise of control*. New York.
- Bar-On, R./Tranel, D./Denburg, N.L. and Bechara, A. (2003): Exploring the neurological substrate of emotional and social intelligence. *Brain* 126 (8), 1790-1800.
- Bechara, A./Tranel, D. and Damasio, A.R. (2000): Poor judgement in spite of high intellect. In: Bar-On, R. and Parker, J.D.A. (Eds.): *The handbook of emotional intelligence*. San Francisco.
- Chen, X. and French, D.C. (2008): Children's social competence in cultural context. *Annual Review of Psychology* 59, 591-616.
- Costa, P.T. and McCrae, R.R. (1992): *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) professional manual*. Odessa.
- Davis, M.H. (1983): Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology* 44, 113-126.
- EU (2005): *Working together for growth and jobs: A new start for the Lisbon Strategy*. European Commission, Communication 33. Brussels.
- Flashman, L.A./Andreasen, N.C./Flaum, M. and Swayze, V.W. (1998): Intelligence and regional brain volumes in normal controls. *Intelligence* 25 (3), 149-160.
- Gardner, H. (1999): *Intelligence Reframed. Multiple intelligences for the 21st century*. New York.
- Goleman, D. (1995): *Emotional intelligence*. New York.
- Grigorenko, E.L. (2000): Heritability and intelligence. In: Sternberg, R.J. (Ed.): *Handbook of Intelligence*. Cambridge.
- Goodman, R. (2001): Psychometric properties of the strengths and difficulties questionnaire (SDQ). *Journal of the American Academy of Child & Adolescent Psychiatry* 40 (11), 1337-1345.
- Gosling, S.D./Rentfrow, P.J. and Swann, W.B. Jr. (2003): A Very Brief Measure of the Big Five Personality Domains. *Journal of Research in Personality* 37 (6), 504-528.
- Hagerty, M.R./Cummins, R.A./Ferriss, A.L./Michalos, K.L.A.C./Peterson, M. et al. (2007): Quality of life indices for national policy: Review and agenda for research. *Social Indicators Research* 55 (1), 1-96.
- Halpern, D. (2005): *Social Capital*. Cambridge.
- Inglehart, R. (1997): *Modernization and postmodernization. Cultural, economic, and political change in 43 societies*. Princeton.

- Kendall, Ph.C. and Wilcox, L.E. (1979): Self-control in children: Development of a rating scale *Journal of Consulting and Clinical Psychology* 47 (6), 1020-1029.
- Kihlstrom, J.F. and Cantor, N. (2000): Social intelligence. In: Sternberg, R.J. (Ed.): *Handbook of intelligence*. Cambridge.
- McCrae, R.R. and Costa, P.T. (2004): *A contemplated revision of the NEO Five-Factor Inventory*. *Personality and Individual Differences* 36 (3), 587-596.
- Putnam, R.D. (2000): *Bowling alone: the collapse and revival of American community*. New York.
- Rampus, K. (1947): Social Competence. *Journal of Abnormal and Social Psychology* 26, 681-687.
- Rosenberg, M. (1979): *Conceiving the self*. New York.
- Rotter, J.B. (1990): Internal versus external control reinforcement. *American Psychologist* 45 (4), 489-493.
- Rydell, A.-M./Hagekull, B. and Bohlin, G. (1997): Measurement of two social competence aspects in childhood. *Developmental Psychology* 33 (5), 824-833.
- Saarni, C. (2000): Emotional competence: A developmental perspective. In: Bar-On, R. and Parker, J.D.A. (Eds.): *The handbook of emotional intelligence*. San Francisco.
- Sarason, B.R. (1981): The dimensions of social competence: Contributions from a variety of research areas. In: Wine, J.D. and Smye, M.D. (Eds.): *Social competence*. New York.
- Sarason, B.R. (Ed.) (1981): *The dimensions of social competence: Contributions from a variety of research areas*. New York.
- Schwarzer, R. (1993): *Measurement of perceived self-efficacy: Psychometric scales for cross-cultural research*. Berlin.
- Tajfel, H. (1981): *Human groups and social categories*. Cambridge.
- Walker, R.E. and Foley, J.M. (1973): Social intelligence: its history and measurement. *Psychological Reports* 33, 839-864.
- Waters, E. and Sroufe, A.L. (1983): Social competence as a developmental construct. *Developmental Review* 3 (1), 79-97.



## 2.5 Subjective Indicators

*Beatrice Rammstedt*

Contact:

Beatrice Rammstedt  
GESIS – Leibniz-Institut für Sozialwissenschaften  
P.O. Box 12 21 55  
68072 Mannheim  
Germany  
e-mail: [beatrice.rammstedt\[at\]gesis.org](mailto:beatrice.rammstedt[at]gesis.org)



## **Abstract**

Subjective indicators have been proven to possess predictive power for a large array of social and economic outcomes. However, most of these measures face serious psychometric shortcomings, namely that the items used are not psychometrically investigated. Furthermore, various different item phrasing and response formats are used in different surveys for the assessment of one and the same construct. The present paper makes several recommendations to improve the quality and thus also the acceptance and usage of subjective indicators. These include the development of more ultra-short but multi-item measures for subjective indicators. In addition, surveys should try to use the same form of measurement (i.e., the same item phrasings and the same response scales). In terms of psychometric properties, the report recommends that the reliability and validity of the indicators be investigated in as much depth as possible. In addition, suggestions are made regarding how best to investigate the respondent's judgmental process for the measurement of subjective indicators, which will allow researchers to obtain a clearer picture of how the item is understood by the respondent and the cues on which he bases his judgment.

Keywords: subjective indicators, reliability, validity, multi-item instruments, cognitive interviews

## **1. Introduction**

Certainly it is possible to assess a person's state of health by investigating patient records, counting the number of sick days, or noting the amount of pharmaceuticals the person takes. However, these are only proxies for the actual state of health. What is more important is the perceived, that is, the subjective state of health. A person going to work even though he feels sick, who is not taking medication, and who does not visit a doctor will be much less productive than a person who perceives himself as healthy.

## **2. Definition: What are subjective indicators?**

At first sight it might be assumed that the distinction between objective and subjective indicators is relatively unambiguous. Objective measurement is based on explicit criteria and is performed by external observers (e.g., Veenhofen 2007). Taking again the example of the individual's state of health, it can be assessed by objective indicators such as antigens in the blood or by the income level; it can be assessed by the annual pay check. However, such objective criteria may also be assessed by subjective mea-

asures, for instance, by use of self-ratings. Are such self-ratings (thus subjectively assessed measures of health or income) still objective or do they become subjective measures? Alternately, one could think of subjective criteria that is assessed by objective measures – such as individually perceived insecurity recorded by the installation of alarm systems or the training in self-defense. Would this be an objective or a subjective indicator? We thus need to differentiate between objective and subjective in both cases, the side of the criterion or substance that is being assessed as well as the assessment itself. Both can be either objective or subjective.

Subjective indicators are often defined as information that includes some kind of a subjective component, such as personal perception or a personal evaluation (e.g., Noll 2001). This definition focuses exclusively on the subjective measurement of any type of criterion. Following the above-described differentiation, this definition includes subjective measures of both objective and subjective criteria. In light of new thinking on this subject, however, this definition is far too broad because it includes any self-reported data and thus can in no way be differentiated by criterion. In the following report, subjective indicators are defined as subjective measures of a subjective criterion. Thus, a comparably strong definition is chosen, which, however, will make it possible to narrow the broad and often somewhat vague field of subjective indicators. Subjective indicators – in this sense – are defined as *subjective information about a subjective criterion*. Subjective indicators thus include constructs like satisfaction, worries, or trust.

### **3. The reason for assessing subjective indicators**

Why do we assess subjective indicators at all, although we know that subjective measures like self-ratings have strong shortcomings since they are – as the label says – subjective. To answer this question we need to disentangle the subjective indicator once again in the subjective substance and its subjective measurement.

Why are researchers interested in subjective substances such as happiness, worries, or values? A common objection has it that such matters are unstable over time, that they are incomparable across cultures and even individuals, that they are unintelligible since they are implicit and therefore their “true score” can hardly be investigated (e.g., Veenhoven 2007). However, research findings have shown that these objections do not apply to *all* subjective criteria. There seem to be several subjective constructs, such as overall life satisfaction, of which individuals have a clear, stable, and comparable understanding. Therefore, even though the rating itself is subjective, the construct it is based on seems to be comparable across individuals. Fur-

thermore, these subjective criteria often have a high impact on external and objective criteria, such as objective welfare or suicide rate. Thus, for the prediction of several important life or societal outcomes, the measurement of subjective substances is indispensable.

However, the question might also be raised, why are these subjective substances measured by self-reports rather than using more objective data, such as indicators for quality of life conditions, to estimate well-being? One reason is that it is often much more time-consuming to gather information on such objective measures than simply to ask the respondent about his or her well-being. The most important reason, however, is that such objective indices do not seem to be an appropriate proxy for the individually perceived criterion. The individual's "true score" in the subjective criterion is only partially based on objective measures, such as, for example, life circumstances. The other part of this "true score" is based on multiple aspects including the individual's fit within his or her environment and the capacity to cope with these life circumstances. This latter part, however, can hardly be assessed by objective proxies.

In sum, research has shown that subjective criteria are fruitful and valid predictors for an array of social and economic outcomes. These subjective criteria are best assessed by means of subjective measures and thus in the form of subjective indicators.

#### **4. Typical subjective indicators in social and economic surveys**

In social and economic surveys there are a number of subjective indicators which are assessed relatively regularly. Instead of providing a complete list, we will give a short overview of the most frequently assessed subjective indicators.

The most prominent criterion for a subjective indicator is *satisfaction* or *subjective well-being*. This can be either operationalized as satisfaction with life in general or as satisfaction with specific aspects of one's life, such as work, marriage, or living conditions. Satisfaction is regularly assessed in numerous studies, such as the German Socio-Economic Panel (SOEP, *Sozio-oekonomisches Panel*), the European Social Survey (ESS), the German Welfare Survey, the European Values Study, and the World Values Survey (WVS).

The counterpart to satisfaction is *worries*. Worries are often assessed in the field of job security or health. Thus, worries are usually assessed based on specific aspects of the respondent's life. Studies that regularly assess worries are again the SOEP and the German Welfare Study.

Another frequently measured complex of subjective indicators in social surveys are measures of *trust*, such as trust in government, in democracy, etc., or even general trust in other people. Various questions regarding trust have been asked in the German General Social Survey (ALLBUS, *Allgemeine Bevölkerungsumfrage der Sozialwissenschaften*), in the Eurobarometer, in the International Social Survey Programme (ISSP), and in the WVS.

## **5. The assessment of the most prominent subjective indicators**

Usually satisfaction, worries, and trust are measured by single item indicators. Typical phrasings of these questions are given in table 1. Respondents are asked to rate their degree of satisfaction, worries, or trust as outlined in the given item formulation on a Likert-scale ranging from “not at all [satisfied, worried, trusting]” to “fully [satisfied, worried, trusting].”

Research, however, has shown that psychometric properties of multi-item scales, as used in psychology for example, are significantly higher than for single-item measures. In multi-item scales, item responses are aggregated to a sum or mean score of the underlying construct. In general, the broader the construct of interest is (e.g., general life satisfaction vs. satisfaction with income), the more multi-item scales should be used. Broader constructs cannot be validly assessed by only one indicator.

Surveys, however, face tight time constraints. Therefore, the number of items used for multi-item scales needs to be restricted. In recent years, efforts have been made to construct ultra-short measures for constructs like personality or values to obtain multiple-item measures that can be used under the tight time restrictions of surveys. This approach could also be very fruitful for subjective indicators. On the one hand it would be a way to reduce lengthy item batteries, on the other broad constructs could be assessed with greater validity.

## 6. How can subjective indicators be validated?

As mentioned above, subjective measures of subjective substances are hardly testable against their supposed “true scores.” The ratings represent feelings, evaluations, or conditions within the respondent that can hardly be perceived from an outside observer. Thus, the respondent’s “true score” remains a mystery.

There are, however, several means available to investigate the quality of a self-report and to gain insight into the process used in making the judgment. If the individual’s response to a subjective indicator is reliable, respondents should answer the same question in the same way if asked again. To avoid memory effects when responding the second time, a substantial time interval should be established between the two administrations, since it can be assumed that a judgment about one’s overall life satisfaction or job security is relatively stable over time, say across several weeks (given that no serious life event happens in that period of time). Therefore, one way of investigating the quality of a subjective indicator is to investigate its stability by administering the same question to the same set of respondents twice. A measurement can be regarded as reliable if the retest stability, thus the correlation between the two administrations, is high.

Another way of investigating the reliability of a subjective indicator is to use multiple items for its assessment. Since all items are assumed to measure the same underlying construct, it can be assumed that all items are positively related to each other. The standardized intercorrelation among this set of items (Cronbach Alpha coefficient) could therefore be regarded as a measure of the measurement reliability.

Testing the validity of the measurement of a subjective indicator poses a larger challenge. Validity is usually investigated by comparing self-estimates to peer or expert ratings or by comparing it to an external (objective) criterion (e.g., self-rated health to results of bio-chemical tests or by the number of doctor’s consultations). Such comparisons are difficult, if not impossible, if the substance rated is subjective and not objectively measurable and perceivable by others – as in the case of health, for example. The only way of shedding some light onto the validity of subjective indicators would be to investigate their construct validity. This could either be done – in the case of multi-item indicators – by a factor analytic approach to test, if all items assumed to measure one indicator from a common factor; or, in a multitrait-multimethod design, by testing several subjective indicators, for example, assessed by self-ratings and peer ratings.<sup>1</sup> Alternatively, this construct validation could be more theoretically driven by investigating

---

1 Peer ratings are also perceived as subjective indicators because they are a subjective judgment about a target person.

hypothesized external correlations. For example, it could be assumed that a person highly worried about his financial situation would be less likely to put high amounts of money in high-risk investments. This form of validation depends significantly on the creation of plausible hypotheses. To develop such hypotheses we first of all need to learn more about the respondent's process of understanding the question and – most importantly – about the respondent's process of developing the judgment that is then given as an answer to the item.

One way of investigating respondent judgmental processes that has been very prominent in recent studies is the qualitative approach of cognitive interviews. In these cognitive interviews, a small number of respondents (preferably rather heterogeneous with regard to their socio-demographic background) are asked to respond to the items in question. In doing so they are either invited to think aloud or to paraphrase specific terms or to explain why they chose a specific answer category. This approach can significantly increase our knowledge about the respondent's understanding, interpretation, and response to a subjective indicator item. This becomes even more valuable the less concrete an item is. Asking a respondent to think aloud while answering the item "Do you like to go to parties" – a typical extraversion item – would probably not give us as much insight into the respondents mind as letting him think aloud about whether he is satisfied with life in general. This latter investigation could help inform us about cues like health, partnership, children, a secure job, or the financial situation, that a person has used to perceive himself or herself as either satisfied or dissatisfied. These cues can then be used to develop hypotheses for assumable correlations. For example, if all respondents refer first to their health as a cue for their level of satisfaction, a significant positive relation between these two indicators can be assumed and tested in terms of validity. Health, however, will never be the only cue and cannot be regarded as a proxy for life satisfaction. Thus, a substantial but not maximal correlation can be anticipated.

Parallel to the development of cognitive interviews, an alternative quantitative approach to shedding light on the individual judgmental process has also been developed (Jasso 2006). In this factorial survey method, respondents are asked to rate the level of a specified outcome variable (e.g., well-being or healthiness) based on a given fictitious description of a person and its characteristics (age, gender, income, eating habits, housing, etc.). Respondents will be rating a large set of such fictitious descriptions, termed "vignettes." Based on these ratings, the implicitly used equations for assigning outcomes like well-being or healthiness can be retrieved using statistical techniques.

In sum, the quality of subjective indicators can be investigated by using methods such as test-retest reliability, Cronbach Alpha coefficients, construct

validity in form of factorial structure or hypothesized correlations with external criteria and by use of cognitive interviews and vignette techniques.

## **7. Recommendations**

Subjective indicators have been proven to possess predictive power for a large array of social and economic outcomes. For most of them, only single item measures are used. The phrasing of items and the response formats differ substantially across different surveys. The items used are in most cases not tested with regard to their psychometric properties. This, however, can be done with very simple means. We therefore recommend the following:

- Subjective indicators should become more widely accepted and investigated as they are proven to possess a substantial predictive power.
- More ultra-short scale measures should be developed and validated
- In order to reach more comparability across studies and thus more comparable results, and in order to profit most from validated measures, surveys should try to use the same form of measurement, (i.e., the same item phrasings and the same response scales).
- Even though the validation of subjective indicators cannot as easily be conducted as it is for objective ones, we strongly recommend that the reliability and validity of these indicators be investigated in as much depth as possible.
- Furthermore, we suggest investigating the respondent's judgmental process for subjective indicator measures by the use of cognitive interviews and/or by the vignette technique. This makes it possible to obtain a clearer picture of how the item is understood by the respondent and on which cues he bases his judgment.
- If results of cognitive interviews indicate that individuals strongly vary in their understanding of and in their way of responding to these single-item indicators, researchers should examine whether it would be more fruitful and thus more valid to use multiple indicators that could be, in turn, less abstract.

Table 1. Typical item phrasings for measures of satisfaction, worries, and trust.

General life satisfaction (WVS)	All things considered, how satisfied are you with your life as a whole these days? Using this card on which 1 means you are "completely dissatisfied" and 10 means you are "completely satisfied" where would you put your satisfaction with your life as a whole?
Domain specific satisfaction (GSOEP)	How satisfied are you today with the following areas of your life? How satisfied are you with... your health your sleep your job your housework your household income your schooling and professional education your place of dwelling your free time your family life the child care available?
Domain specific worries (GSOEP)	What is your attitude towards the following areas – are you concerned about them? General economic development Your own economic situation Your health Environmental protection Maintaining peace Global terrorism Crime in Germany Consequences of the expanding of the EU to the east Immigration to Germany Hostility towards foreigners or minorities in Germany <i>If you are employed:</i> Your job security?
Trust in others (WVS)	Generally speaking, would you say that most people can be trusted or that you need to be very careful in dealing with people?  I'd like to ask you how much you trust people from various groups. Could you tell me for each whether you trust people from this group completely, somewhat, not very much or not at all? - your family - your neighbourhood



	<ul style="list-style-type: none"> <li>- people you know personally</li> <li>- people you meet for the first time</li> <li>- people of another religion</li> <li>- people of another nationality</li> </ul>
Trust in institutions (ALLBUS)	<p>I am now going to read out a number of public institutions and organisations. Please tell me for each institution or organisation how much trust you place in it.</p> <p>Please use this scale.</p> <p>1 means you have absolutely no trust at all  7 means you have a great deal of trust</p> <p>You can differentiate your answers using the numbers in between. What about the –</p> <p>Health service  German constitutional court  German Parliament  Municipal administration  Army  Catholic church  Protestant church  Judicial system  Television  Newspapers  Universities and other institutes of higher education  German government  Trade unions  Police  Job centres  State pension system  Employer associations?</p>

#### References:

- Jasso, G. (2006): Factorial survey methods for studying beliefs and judgments. *Sociological Methods and Research* 34, 334-423.
- Noll, H.-H. (2001): Subjektive Indikatoren: Expertise für die Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik. In: Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik (KVI) (Ed.): *Wege zu einer besseren informationellen Infrastruktur*. Baden-Baden.
- Veenhoven, R. (2007): Subjective measures of well-being. In: McGillivray, M. (Ed.): *Human Well-Being – Concept and Measurement*. Houndsmills.

### 3. EDUCATION AND RESEARCH

#### 3.1 Education Across the Life Course

*Hans-Peter Blossfeld*

Contact:

Hans-Peter Blossfeld  
University of Bamberg  
Chair of Sociology I  
Faculty of Social Sciences, Business Administration and Economics  
PO Box 1549  
Lichtenhaidestr. 11  
96045 Bamberg  
Germany  
e-mail: Soziologie1[at]uni-bamberg.de

## Abstract

There is enormous demand in Germany for high-quality longitudinal research on education. In particular, there is a clear need for both analytical and methodological progress in order to understand educational pathways through the life course and how they lead to different outcomes. This paper identifies the theoretical and methodological challenges of studying education across the life course and describes the structure of the National Educational Panel Study (NEPS) in Germany.

**Keywords:** competence development, educational decisions, formal, informal and nonformal educational environments, returns to education, educational trajectories, life-course research, longitudinal analysis, panel data

## 1. Research questions

Germany, like other modern industrialized societies, has evolved into a knowledge-based economy. More than in the past, today education is a life-long process in which individuals continually learn in formal, non-formal, and informal environments. Individuals' educational careers and competencies and how they unfold over the life course in relation to family, educational institutions, workplaces, and private life are therefore a topic of considerable national interest. There are several social, economic and demographic changes that make education particularly important in modern societies. Some examples include:

- The dramatic decline in the number of unskilled jobs that has taken place over the course of technological change raises issues concerning the proportion of youth who leave the educational system with poor competencies and/or without a certificate. Questions thus arise, such as which factors influence educational participation and achievement in educational institutions? How are basic competencies related to school success, certificates, the transition into the labor market, and job careers?
- Technological change has also led to an increase in the number of service, professional, and engineering positions that require a range of social, communicational, and problem-solving skills. This upgrading of the job structure has enhanced the value of education, science competencies, and soft skills on the labor market and in society as a whole. Such changes inevitably affect the content and processes of learning in schools, vocational training systems, and universities, triggering questions such as what do students learn and which kinds of competencies do individuals need in a knowledge society? How can we engender motivation and interest in

reading, mathematics, and science? How is the development of reading, mathematical, and science competencies affected over the life course by different learning environments such as school, work, home, and community? How can we quantify the economic and social benefits of these competencies?

- New information and media technologies continue to have a growing impact on our daily lives and work. They demand the ability to communicate and share information. This raises questions about how well the educational system provides the relevant training for ICT (Information and Communication Technologies) literacy, social competencies, and personal skills to meet these demands.
- Increasing worldwide competition and globalization are stepping up the pace of social and economic change, making it necessary for individuals to exhibit greater flexibility and adaptability both at their workplace and in society. The ability to learn new skills and to adapt has become an important requirement for securing jobs and being a successful citizen. Therefore, the question is, how does the initial full-time educational investment influence these learning skills and attitudes toward educational institutions, further education, and learning situations over the life course? How can we improve self-regulated learning so that individuals realize opportunities to take on challenging tasks, practice their learning, and develop a deep understanding of their subject matter?
- Germany is also undergoing considerable demographic changes: there is a decline in fertility, an aging population, an increasing diversity of ethnic background in the population, and a rising level of instability in family life. All these changes have direct consequences for individual educational processes (e.g., the impact of parental divorce on the educational careers of children and their competence development) and for educational institutions as a whole (e.g., when *Hauptschule* and *Realschule* are combined because of the declining number of students, or when older workers have to increase their participation in education during late adulthood). The capacity to follow individuals through the life course and to observe how educational experiences, competencies, and behavior are influenced by the formal, non-formal, or informal contexts in which they find themselves grants longitudinal studies a major role in understanding the role of education in modern society.
- Empirical studies such as PISA (Programme for International Student Assessment) demonstrate that the quality and quantity of schooling that individuals acquire still depend to a large extent on the advantages or

disadvantages that parents confer throughout childhood, adolescence, and adulthood. However, most of these educational inequalities are only documented on the basis of cross-sectional data. The causal mechanisms that produce primary and secondary outcomes as well as the cumulative processes over the educational career are still subject to considerable controversy in the literature on educational inequality and can only be studied reasonably on the basis of longitudinal data. In particular, the complex and subtle role that schools, vocational training institutions, universities, and continuing education play in the maintenance of inequality and the allocation of persons to unequal positions in Germany across generations and within the life course is still not well understood.

## 2. Status quo: Databases and access

National and international school performance studies or student assessment studies such as TIMSS,<sup>1</sup> PISA, PIRLS,<sup>2</sup> or DESI<sup>3</sup> have developed competence tests in different domains (mainly in the domains of reading, math, and science literacy or in English as a second language). These studies provide information on the distribution of competencies within the Federal Republic of Germany in comparison with other countries as a function of school type attended, social background, and student gender. However, they are cross-sectional and therefore provide only a snapshot of different students at a particular point in their educational careers. Successive snapshots in a series of cross-sectional surveys highlight changes in the structure as a whole. Yet, they do not show the changing (and sometimes unchanging) experiences of individual students as their educational careers progress.

Several longitudinal studies have already been carried out in Germany that broaden the knowledge derived from these cross-sectional studies by providing more information about the causes of established competence development and educational decisions. The available longitudinal studies can be assigned to the following four areas: (1) childhood development, (2) transitions and competence developments in elementary and secondary school, (3) transitions from school to vocational training and university, and (4) life-course research with a strong emphasis on educational and employment careers and family-related processes:

---

1 Trends in International Mathematics and Science Study.

2 Progress in International Reading Literacy Study.

3 Assessment of Student Achievements in German and English as a Foreign Language (DESI, *Deutsch-Englisch-Schülerleistungen-International*).

- The studies concerning childhood development are national (DJI<sup>4</sup> Children's Panel) or regional longitudinal studies (BiKS,<sup>5</sup> LOGIK<sup>6</sup>) on competence and personality development in children and on the transition from kindergarten to elementary school.
- The majority of longitudinal studies were carried out on educational development within schools. Among these regionally designed longitudinal studies, we can differentiate two types: The first concentrates on competence development within one level of education (SCHOLASTIK,<sup>7</sup> BeLesen,<sup>8</sup> and Hannoversche Grundschulstudie in elementary school; PALMA<sup>9</sup> in the lower secondary school), whereas the second type predominantly examines transitions between two stages of education (BiKS in Hesse and Bavaria, KESS<sup>10</sup> in Hamburg, Koala-S in Bavaria and Saxony). However, some studies have a strong focus on competence development as well as on transitions (ELEMENT<sup>11</sup> in Berlin, BIJU<sup>12</sup> in Mecklenburg-Western Pomerania, Saxony-Anhalt, North Rhine-Westphalia, and Berlin). Only two nationwide studies have a rudimentary longitudinal character: TIMSS 1995 and PISA 2003. TIMSS 1995 has tested students in the 7th grade (1994) and then again one year later (1995). PISA 2003 has been expanded by a second wave (PISA-I-plus). Ninth graders from the intermediate and academically oriented track were tested one year later in tenth grade in order to analyze how they had progressed in mathematics and sciences and what the determining factors were.

---

4 German Youth Institute (DJI, *Deutsches Jugendinstitut*).

5 Educational Processes, Competence Development, and Selection Decisions in Pre- and Primary School Age (BiKS, *Bildungsprozesse, Kompetenzentwicklungen und Selektionsentscheidungen im Vor- und Grundschulalter*).

6 German Longitudinal Study on the Genesis of Individual Competencies (LOGIK, *Longitudinalstudie zur Genese individueller Kompetenzen*).

7 Organized Learning Opportunities at School and the Socialization of Talents, Interests, and Competencies (SCHOLASTIK, *Schulorganisierte Lernangebote und die Sozialisation von Talenten, Interessen und Kompetenzen*).

8 Berlin Longitudinal Study on the Literary Development of Primary School Children (BeLesen, *Berliner Längsschnittstudie zur Lesekompetenzentwicklung von Grundschulkindern*).

9 Project on the Analysis of Performance Development in Mathematics (PALMA, *Projekt zur Analyse der Leistungsentwicklung in Mathematik*).

10 Competencies and Attitudes of School Students (KESS, *Kompetenzen und Einstellungen von Schülerinnen und Schülern*).

11 Study on Competencies in Reading and Mathematics (ELEMENT, *Erhebung zum Lese- und Mathematikverständnis*).

12 Learning Processes, Educational Careers, and Psychosocial Development in Adolescence and Young Adulthood (BIJU, *Bildungsverläufe und psychosoziale Entwicklung im Jugend- und jungen Erwachsenenalter*).



- For participation in the university and the labor market entries of academics, the Higher Education Information System (HIS, *Hochschul-Informationen-System*) has conducted national longitudinal studies in which, however, no performance-based competence measurements were included. One of the HIS panels covers a cohort of secondary school graduates qualified for higher education, follows their transition into the university or vocational training programs and their subsequent educational career for a period of three and a half years after leaving school. The HIS survey of graduates concentrates on the transition from university to the labor market and into the professional career. The DJI Transition Panel, implemented by the German Youth Institute, focuses on the transition of “disadvantaged” students who have finished the lower school track, and follows their paths into the vocational training system and their entry into the labor market (no competence tests have been conducted). There is also the ULME<sup>13</sup> study in Hamburg that is testing competence development from entry until the end of a course of study at vocational schools, independent of whether the vocational school is full- or part-time.
- The following longitudinal studies differ from those previously summarized. Their focus is on a longer time span. The study of former students from academically oriented secondary schools (*Gymnasium*) follows student careers (beginning in the 10th grade) and examines college and professional education as well as gainful employment in North Rhine-Westphalia over a period of twenty-eight years. In addition, the GLHS (German Life History Study) collected data retrospectively from several birth cohorts on their previous educational and employment career as well as their family history in Germany. No competence tests were included in the GLHS. Since the beginning of the 1990s, individual biographies of East Germans have been surveyed in order to obtain detailed information on lives before, during, and after reunification. Finally, the German Socio-Economic Panel (SOEP, *Sozio-oekonomisches Panel*), a general public survey carried out every year in Germany since 1984, includes large samples of West and East Germans as well as various groups of immigrants. The SOEP combines retrospective data on the work and family-related event history with prospective panel data on, among others, job and income mobility, educational participation, family status, and life satisfaction in different domains.

---

13 Study on Achievement, Motivation, and Attitudes of Students at the Beginning of Vocational Education (ULME, *Untersuchungen der Leistungen, Motivation und Einstellungen zu Beginn der beruflichen Ausbildung*).

This short overview of available longitudinal studies conducted in Germany reveals that there is only one genuine nationwide panel study, the SOEP. However, this study includes no detailed data on changes in educational contexts and no information on the development of competencies. Based on the other longitudinal studies that do measure competencies, only limited conclusions can be drawn. This is because they either confine themselves to a certain region within Germany or concentrate primarily on one stage of education or a specific transition in the educational career. These studies make it impossible to understand how the competencies of individuals develop over the life course, how they interact with educational decisions at various critical transitions in their careers, and how competencies are influenced by the family and the arrangements of teaching and learning processes in kindergarten, school, professional education, and university. It is also unclear how competencies are related to the achievement of educational qualifications, and which competencies are responsible for labor market success and a successful private and social life.

Other European and North American countries have a longer tradition in conducting educational panel studies that include the assessment of competencies, skills, or intelligence components. Kristen et al. (2005) provide an extensive overview of studies conducted in Canada, France, the Netherlands, Sweden, the United Kingdom, and the United States. In these countries, different approaches have been chosen to obtain longitudinal information on education. These are mainly either long-running cohort studies that collect data on an individual's life over a long period or short-term studies on a specific stage of the educational career.

In the United Kingdom, large birth cohort studies have been carried out with educational topics. These studies started in 1958 with the National Child Development Study (NCDS), continuing with the British Cohort Study (BCS70) in 1970 and the Millennium Cohort Study (MCS) in 2000-2001. While in the 1958 NCDS, the distances between the panel points fluctuated between four and ten years, the Millennium Cohort Study intends to have a much smaller time span.

The Effective Provision of Pre-School Education (EPPE) project running from 1997 to 2003 is the first major study in the United Kingdom to focus specifically on the effectiveness of early childhood education (the EPPE 3-11 Project from the years 2003-2008 builds on the original EPPE study). The EPPE project is thus a large-scale longitudinal study of the progress and development of children in and from various types of preschool education (from ages three to eleven).

England and Wales used a research strategy to focus on a short relevant sequence of the educational career. The Youth Cohort Study (YCS) is a repeated short panel study providing insight into the transition from secondary school to further education and to the labor market. Another variant of a

cohort study focusing on careers after completing compulsory education is the Canadian Youth in Transition Survey (YITS). This survey attempts to follow students over a longer period of time. The study started with two subsamples in 2000. The members of the first sample were 15 years old, those of the second aged from 18 to 20. For the younger cohorts, only competencies were tested within the framework of PISA 2000. A similar strategy has been implemented by the Swiss Transitions from Education to Employment longitudinal study (TREE), which also started in the year 2000 and measured competencies within the PISA framework only for this first year. Recently (in 2006), the Swiss Survey of Children and Youth COCON (competence and context), was also initiated. This study investigates the social conditions, life experiences, and psychosocial development of children and youth in the German- and French-speaking parts of Switzerland from a life-course perspective. The longitudinal part of this study follows up two cohorts: six-year-olds (middle childhood) and fifteen-year-olds (middle adolescence). In the United States, there have been a large number of different longitudinal education studies. Their main goal has been to analyze the educational, professional, and personal development at different points in the educational career and to identify the role played by personal, family, social, institutional, and cultural factors (NCES 2003). Most of these cohort studies have four or five observation points, and begin at Grade 10 or 12. These are high school studies focusing on the transition to postsecondary education and on labor market entries, including the National Longitudinal Study of the High School Class of 1972 (NLS-72), the National Education Longitudinal Study of 1988 (NELS-88), the Educational Longitudinal Study (ELS 2002), and High School and Beyond (HS&B). Some studies concentrate on students in tertiary education and their labor market entry, such as the Beginning Postsecondary Students Longitudinal Study (BPS), and Baccalaureate & Beyond (B&B). In recent years, as part of the Early Childhood Longitudinal Study (ECLS), two cohort studies have commenced that focus on development at early ages. One cohort starts with the newborn, the other with children attending kindergarten or preschool institutions. The situation for education data in the US can be described as an additive, repeated cohort sampling. This means that different cohorts are available for children and students at a specific stage in the educational career. The complete data pattern of these cohorts delivers sequences reaching from birth up to age thirty. However, there still remains a gap at the lower secondary level. Finally, there are international longitudinal studies on school-to-work transitions. Most of these longitudinal studies focus on educational biographies and transitions, some of them have conducted cross-sectional competence and/or skill measures (in different domains, mostly by including the youths tested in the PISA Studies), but only few have started to survey longitudinal competence developments.

This brief overview of different longitudinal studies using different longitudinal designs makes it clear that birth cohort studies take too much time to acquire a “complete” picture of the educational career. To study children’s development and transitions until the end of the secondary school level would take nearly twenty years. Therefore, it is more efficient to concentrate on important sequences in the educational career. Samples must be drawn for every relevant sequence. Such a multicohort sequence design quickly provides the relevant information. In order to capture the influence of educational reforms and social change, new cohorts have to be sampled repeatedly. Such a strategy is comparable to the one followed by the US-National Center for Education Statistics.

In summary, there is enormous demand for high-quality longitudinal educational research in Germany. In particular, there is a clear need for both analytical and methodological progress in order to understand educational pathways through the life course and how they lead to different outcomes.

### **3. Future developments in Germany: The National Educational Panel Study**

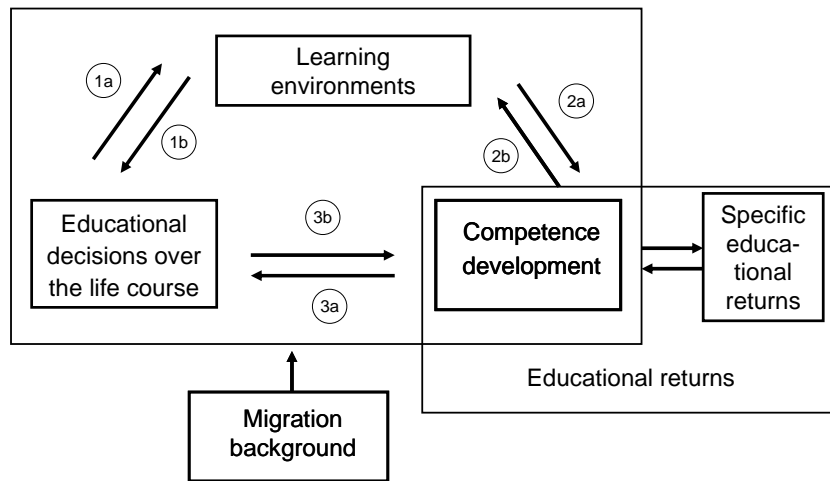
The National Educational Panel Study (NEPS), supported by funds from the Federal Ministry of Education and Research (BMBF, *Bundesministerium für Bildung und Forschung*), started in August 2008. The basic design and organization of the NEPS can be summarized as follows.

#### *3.1 Theoretical framework of the NEPS*

The key theoretical assumptions of the NEPS as an instrument for studying education over the life course can be best summarized in a diagram. Figure 1 shows that individuals’ educational trajectories over the life course are the result of a dynamic system, creating a complex, time-related interdependence of educational decision making, educational processes within different learning environments, and competence development: (1a) Decisions (by parents, students, or teachers) determine whether and to what extent individuals participate in educationally relevant social and institutional contexts; (1b) participation in formal, non-formal, and informal learning environments, in turn, will influence further educational decision making; (2a) educational processes within learning environments are supposed to have an effect on competence development; (2b) competence development, in turn, will influence future opportunities to participate in social and institutional contexts; (3a) competence development will also affect the processes of educational

decision making; and (3b) educational decisions will influence the future competence development over the life course. Focusing on these three key theoretical dimensions and their time-dependent interaction mechanisms, which generate change and development over the life course, establishes a foundation for powerful explanations and evidence-based research in the NEPS.

Figure 1: Dynamic Interdependence of Educational Decision Making, Participation in Learning Environments, and Competence Development Over the Life Course



It is well known from several recent studies that the educational outcomes of immigrants' children differ substantially from those of their peers from native families. These differences are likely to exist across the whole life course and are conditioned by several specific theoretical mechanisms. In addition to the three main theoretical dimensions, a fourth theoretical dimension of the NEPS is therefore concerned with the educational career of immigrants and their descendants. In order to account for ethnic inequalities, it is necessary to ask which specific resources and orientations on the level of the individual, the family, the learning environment, and the context (e.g., local community), as well as which institutional (e.g., regulation of transitions, availability of education in German as a second language), societal (e.g., acculturation orientations), and political conditions (e.g., regulations for residency status) impact systematically the success of immigrants and their descendants in the educational system and the labor market. The crucial theoretical and empirical task for the NEPS is therefore to identify the particular mechanisms affecting the competence development and the educational decision processes of immigrants.

Finally, there is a fifth significant theoretical dimension to the NEPS that concerns the issue of educational returns. Given that education unfolds over the life course and myriad interactions and relationships are involved, the measurement and modeling of these returns must concentrate on the changes of these outcomes over the life course and the complex dynamic interaction processes that take place when qualifications, competencies, and educational certificates at certain points in the life course are turned into future economic and non-economic returns. Thus, at least three aspects are important for the returns of education over the life course: the first is that education is a lifelong and cumulative process, and that the educational events and experiences in earlier life stages have consequences for later educational processes and competencies. In other words, from a life-course perspective, later educational participation and competence development are themselves returns to earlier educational investments (see figure 1). Second, educational events, considered in causal process models as constituting the causes of economic and non-economic returns, are to a large extent shaped by these returns. Finally, the importance of the different non-economic and economic returns strongly varies over the life course, because they are often associated with certain development stages or are connected to specific life-course transitions. The NEPS will focus on three economic dimensions: (1) reconsidering the effects of education in classic estimates of monetary returns to years or level of education, job opportunities, and job mobility rates; (2) returns to educational reforms; and, most importantly, (3) returns to specific school institutions. Apart from the monetary economic returns and the returns in terms of later education, the NEPS will include additional nonmonetary returns to education in several areas. Nonmarket returns may come most notably in the form of (a) health, (b) family and fertility behavior, (c) reduced crime and deviance, (d) increasing political and social participation, and (e) subjective well-being.

These five theoretical dimensions are called “pillars,” because they will help the consortium to integrate the multicohort sequence design of the NEPS in terms of content, theory, and method and provide a unified mold for the NEPS. In organizational terms, these five theoretical key perspectives are represented by five substantively focused expert groups. Since the NEPS is hosted at Bamberg University, the five pillars will be coordinated and integrated by experts from the Institute for Longitudinal Educational Research Bamberg (INBIL, *Institut für bildungswissenschaftliche Längsschnittforschung*). In addition, two expert groups will support the NEPS: an expert group that will take care of the most important methodological issues of the NEPS, such as sampling design, data cleaning, data archiving, data dissemination, as well as methodological analysis and training; and an expert group on technology-based assessment (TBA) that will support the NEPS with respect to issues involved in computer- and Internet-based assessments.

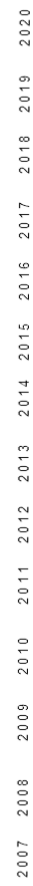
### 3.2 *The multicohort sequence design of the NEPS*

The aim of the NEPS is to provide fast and up-to-date information on educational processes in the various parts of the educational system. Thus, it cannot start with a single birth cohort and then follow it up for 20 years until some of the cohort members eventually leave university. Instead, the consortium has decided to start with several well-chosen cohorts at the same time and to follow these cohorts over longer time spans in their lives. The cohorts will be selected around crucial educational transitions in the German educational system. The multicohort sequence design of the NEPS is shown in figure 2. This design covers: (1) educational processes in kindergarten and the transition to elementary school, (2) educational processes in elementary school and the transition into the tracked secondary school system, (3) processes in the lower secondary school and the transition to upper secondary school, (4) processes of education in upper secondary school and the transition to university or vocational training, (5) educational processes at the university level and the transition of university graduates into the labor market, (6) vocational training and transitions into the labor market, and (7) processes of lifelong learning.

In other words, NEPS will start with five cohorts and follow them up over longer periods in the educational system (Figure 2). It is suggested that four of the five cohort studies begin in the fall of 2010 and then continue with annual observations (the vertical line in figure 2 marks the end of a first 5-year funding period). These five cohorts will be complemented by a panel survey of individuals aged 23 to 64 who have already left full-time education in order to collect data on adult education and lifelong learning. The Institute for Employment Research (IAB, *Institut für Arbeitsmarkt- Und Berufsforschung*) began this survey in 2007 (Figure 2) and it will be integrated into the NEPS in 2009. This multicohort sequence design repeats its structure after some time. However, the next new cohort will not start with kindergarten aged children but with newborn infants (at age six months, see figure 2), therefore allowing not only cohort comparisons but also the identification of age, period, and cohort effects – at least after some time.

In order to develop appropriate instruments for the different educational stages and transitions within the German educational system, the consortium will draw on seven expert groups. The idea behind this is that expert groups on each of the five pillars will cooperate closely with the expert groups responsible for the seven educational stages to develop the necessary questionnaires and competence tests for the NEPS. The representatives of these pillars will ensure that the unified research perspective represented by the five pillars is taken into account by all stage-specific research groups and that the measurement instruments are comparable across the various cohorts of the NEPS.

838





#### **4. Some concluding remarks**

The center of the NEPS consortium is hosted at INBIL at Bamberg University. The role of the INBIL and the division of work between INBIL, the expert groups, and the survey institutes is described in detail in the research proposal (Blossfeld 2008): as a first step, the questionnaires and test instruments for the seven educational stages of the NEPS will be constructed in close collaboration between INBIL and the expert groups. Because the coordinators of the pillars are members of INBIL, it is guaranteed that the multicohort sequence studies will be integrated in terms of concepts, operationalizations, and measurements. After the measurement instruments have been constructed, INBIL will, in a second step, commission survey institutes to carry out the random sampling and the data collection. After data collection, the survey institutes will deliver the datasets to the data center of INBIL. INBIL will then perform data cleaning, validation, coding, scale construction, data weighting, imputation, anonymization, data preparation, data documentation, and so forth. After collection, cleaning, and archiving, the NEPS data will be disseminated to the scientific community as quickly as possible. For this purpose, the NEPS will produce a Scientific Use File and offer training courses on how to effectively exploit the potential of the NEPS for appropriate analyses of all forms.

In sum, the NEPS will establish an excellent scientific evidence base with which to address a broad range of both basic and applied questions in the field of education and to inform policymaking. In particular, the NEPS will provide representative data on the condition of all relevant parts of the educational system in Germany (system monitoring) and offer a better base of scientific evidence for educational reforms and political consultation (system improvement).

#### References:

- Blossfeld, H.-P.(2008): Education as a lifelong process. A proposal for a National Educational Panel Study (NEPS) in Germany. Part A: Overview, Part B: Theories, Operationalizations and Piloting Strategies for the Proposed Measurement, and Part C: Appendices. Faculty of Social and Economic Sciences, Bamberg University.
- Kristen, C./Römmel, A./Müller, W. and Kalter, F. (2005): Longitudinal Studies for Education Reports: European and North American Examples. Report commissioned by the Federal Ministry of Education and Research. Education Reform vol. 10. Berlin. [German version: Längsschnittstudien für die Bildungsberichterstattung – Beispiele aus Europa und Nordamerika. Gutachten im Auftrag des Bundesministeriums für Bildung und Forschung. Bundesministerium für Bildung und Forschung. Bildungsreform Band 10].

## 3.2   Preschool Education

*C. Katharina Spieß*

Contact:

C. Katharina Spieß  
German Institute for Economic Research (DIW Berlin)  
DIW Berlin  
Mohrenstraße 58  
10117 Berlin  
Germany  
e-mail: [kspiess\[at\]diw.de](mailto:kspiess[at]diw.de)

## Abstract

Given the importance of the early stages in a child's life, and taking into account the various initiatives underway to improve preschool programs in Germany, it is remarkable only a few microdatasets cover the field of preschool education in Germany – and even fewer of these are nationally representative datasets. The majority of the existing data provide, at a minimum, basic information on attendance in preschool programs. In principle there are two main groups of data: data comprised of information collected by official statistics and survey data. However, hardly any data is collected that allow researchers to link preschool program information with child outcome data. There is an urgent need for better data on children from age zero to three, as well as for data on children from immigrant families. There is, in particular, a need for good panel data that would permit individual data to be matched with institutional information. Given recent developments in the German data infrastructure, the potential for preschool education research will certainly improve. Nevertheless, there are a number of significant gaps that need to be explicitly addressed. This contribution recommends several key improvements within the field, including better data on the quality of preschool programs, on family context, and on the cost of preschool education. Finally, the paper stresses the need for detailed intervention studies (on a representative, or generalizable level), which can help us to learn more about the most effective and efficient parameters of preschool programs.

Keywords: preschool education, day care, child outcomes

## 1. Research questions

Preschool education refers to education given children before the commencement of statutory education, often between the age of two and compulsory school age. The term “preschool education” generally refers to preschool programs in formal educational settings. The prevalent type of preschool programs in Germany are the “*Kindertageseinrichtungen*,”<sup>1</sup> the generic term that encompasses traditional *Kindergarten*, *Kinderkrippe*, and day care centers, which offer care to children ranging in age from birth to compulsory school age. A broader concept of preschool education might also include family day care (*Tagespflege*), or at least licensed family day care. This is a particularly relevant issue in Germany, where family day care is currently being considered as an alternative to traditional day care programs, at least for younger preschoolers.

---

1 In an even stricter definition the first tier of the German educational system, the so-called *Elementarbereich*, refers to the German *Kindergarten*, which starts at age three. However, a focus solely on the *Elementarbereich* seems too narrow for the analysis of preschool education in Germany today.

Preschool education is embedded in the broader field of early childhood education. This field of educational research views parents and families as integral parts of the early childhood education process, apart from formal educational settings. Thus, for this age group in particular, both the family as a context and the interaction between family and preschool education is of great importance.

Today more than ever, the importance of early childhood education is being recognized and investigated by a range of disciplines, including neuroscience, developmental psychology, educational research, educational economics, and sociology. In all these diverse strands of research there is broad agreement that these early years play a crucial role in the child's development, and are particularly important for his or her later performance.

Considering the importance of this early stage, a great deal of attention should be devoted to preschool programs, particularly if the broader goal is to improve the effectiveness and efficiency of the educational system as a whole. Preschool programs produce a wide range of positive effects. US studies have demonstrated that high-quality programs produce short-term gains in cognitive functioning and longer-term gains in school achievement, including special education placement, high school graduation, and college enrollment. Other positive impacts include better health as adults, reduced criminal activities, and an increase in lifetime earnings. High-quality preschool programs are particularly important for disadvantaged children.<sup>2</sup> Although most of these effects have been demonstrated in the context of model programs, high-quality preschool programs can be considered an effective tool to reduce the achievement gap between poor children and children from more affluent families. In a broader sense, high-quality preschool programs can contribute to increasing an economy's human capital.

Today more than ever before, preschool programs are attracting attention among the German public. Beginning with the "Schröder-government" and increasingly under the first "Merkel-government," initiatives have been launched to augment preschool education for children under the age of three. The Law for Expansion of Daycare (TAG, *Tagesbetreuungsausbaugesetz*) and the Law to Support Children (Kifög, *Kinderförderungsgesetz*) at the federal level represent major steps in this direction. These initiatives have been accompanied by various initiatives at the state and municipal level. The federal states and in particular the municipalities are the agencies actually responsible for the funding of preschool programs. Although these initiatives are motivated by more than just educational objectives,<sup>3</sup> they are critically important for the improvement of preschool education in Germany. The

---

2 For a summary of the results of the relevant US studies, see for instance Karoly et al. (2006).

3 Another important motivation behind these initiatives is the improvement of the reconciliation of work and family life in Germany.

federal government has set a target of providing preschool education or day care to 35 percent of children under the age of three by the year 2013. In addition to these efforts to increase the quantity of preschool in Germany, there are various initiatives to increase preschool quality. In providing these measures, the German government is trying to catch up with the preschool provision rates in other countries such as the Scandinavian countries, France, and Belgium. For instance, in Denmark almost 62 percent, in Norway 44 percent, and in Belgium 34 percent of all children under the age of three attend some kind of formal day care. In Germany the corresponding number was 9 percent (OECD 2006; 2007; UNICEF 2008). It is clear that countries like Sweden and Denmark, widely considered to be the childcare and preschool education leaders in Europe, offer their citizens universal—or nearly universal—high-quality, publicly-funded childcare. However, there are other countries like Germany that have begun to promote increased preschool education attendance. The UK, for instance, has sought to learn from the Nordic childcare model by moving towards a model of universal childcare while focusing on an educational approach for early childhood services.

For older preschoolers (three years and over), the German discussion looks different. In Germany these children have the legal right to a place in a formal day care program, or *Kindertageseinrichtung*; however, at the federal level this legal right only covers four hours per day. The public debate here seems to favor an increase in the number of available slots covering more hours as well as the provision of a lunch. Again, this is an already established standard in some other European countries: in Sweden 63 percent, in Denmark 83 percent, and in France 45 percent of all children aged three to five are enrolled in full-time day care; in Germany the rate was only 29 percent (OECD 2006; 2007).

In terms of the educational aspects of preschool facilities, there is yet another issue that requires attention. Children from families occupying a better socio-economic position are particularly overrepresented in the German preschool system, especially up to age four. Empirical evidence suggests that this is different from other European countries, in particular from the Scandinavian countries (OECD 2006). After the age of four, almost all children in Germany attend preschool (87 percent of all 4–5 year-olds and 91 percent of all 5–6 year-olds).<sup>4</sup> Below the age of four, however, preschool attendance is higher amongst those children with a non-migration background compared to those with a migration background. Nonetheless, preschool attendance at later ages (4–5 and 5–6 year olds) does not differ anymore by migration background. This is significant as preschool education could be an efficient tool for integration.

---

4 See for instance, Konsortium Bildungsberichterstattung (2006: chapter C).

## 2. Status quo: Databases and access

There are only a few microdatasets covering the field of preschool education in Germany – even fewer of these are nationally representative datasets. The majority of these provide, at a minimum, information on attendance in preschool programs. There are two main groups of data: the first includes data collected by official statistics. The most important database of this type is the Children and Youth Services Statistics (KJH-Statistik, *Kinder- und Jugendhilfestatistik*) (see Schilling 2002; Kolvenbach and Taubmann 2006). Since 2006, KJH-Statistik has compiled information on the number of children attending a formal day care program and the number of children attending a publicly funded family day care facility (*Kindertagespflege*). They include information on the staff of these programs and facilities.<sup>5</sup> In addition, data on the age and gender of the children, the country of origin of the parents, the language spoken most at home, and information on special needs for support are also collected. The KJH-Statistik covers the number of hours in care as stipulated in the care contract<sup>6</sup> and the provision of school lunch. These data indicate the type of provider – whether a non-profit provider or for-profit provider. The information on staff details their gender and contracted working time. The information on age, qualifications, occupational status, and field of activity is collected for the majority of the staff members. All data are collected on a yearly basis.<sup>7</sup> Official reporting of these statistics differs by state and district. It is the task of the Office for Children and Youth Services Statistics (AKJ, *Arbeitsstelle Kinder- und Jugendhilfestatistik*) in Dortmund to analyze this data and to promote its use.<sup>8</sup> The microdata can be used via the Research Data Centers of the Statistical Offices of the German *Länder*. Other official statistics no longer cover preschool education: the German Microcensus stopped doing so in 2005. Thus, the Scientific Use Files from the Microcensus provide information on whether the children in the household attend some type of formal day care or preschool only until 2004.

The second set of data is survey data. This group covers the Children's Panel of the German Youth Institute (*DJI, Deutsches Jugendinstitut*).<sup>9</sup> This

---

5 Before 2006 these statistics included information on the number of places available in formal day care programs. This supply-side approach was changed to provide better information on actual attendance rates. For the relevant law, see the Act on the Further Development of Child and Youth Services (KICK, *Kinder- und Jugendhilfeentwicklungsgesetz*), which came into effect in October 2005 (Kolvenbach and Taubmann 2006). Information on family day care facilities was not collected before this date.

6 The daily care hours as contracted do not necessarily correspond to the actual amount of time in day care.

7 Before 2006, they were collected on a four-year basis only.

8 <http://www.akjstat.uni-dortmund.de/>

9 <http://www.dji.de/kinderpanel>. See also Alt (2005).



study was started in 2002 with two cohorts: one of children in their last year of preschool (age five), and one of children in second grade (age eight). In the first wave the sample size was over 2,000 children. Two more waves followed in 2004 and 2005. The DJI Children's Panel covers various topics in addition to information about the preschool education. These include whether a child attends a formal day care program, and information about the provider, actual daily care hours, costs, and the level of parental satisfaction with the preschool program. In addition, information on the child's health and personality is collected as well as data on the family and household. There is a special sample for the Turkish and Russian minorities. Given the panel character of the DJI Children's Panel, this dataset allows for longitudinal research, although the panel covers only three waves. In principle, the DJI study is open to the entire research community.

Another dataset covering preschool education is the German Socio-Economic Panel (SOEP, *Sozio-oekonomisches Panel*).<sup>10</sup> The SOEP is a representative longitudinal study of private households in Germany that has regularly surveyed the same private households, individuals, and families since 1984. Since this time, information on preschool education has also been collected, including data on preschool attendance and daily hours in preschool for all children in the household. Every four years, more information is collected on the type of provider, lunch provision, and parental fees. In 2003, the SOEP started collecting age-specific information as well. In the meantime, special survey instruments have been developed for children in their first year of life, for two- to three-year olds, and for children in their last year of preschool. For all these three age groups, the SOEP collects more detailed information on the child, including his or her preschool education and child outcomes (for more information, see Lohmann et al. 2008: chapters 3.1 and 8.1). Further age-specific survey instruments for children in school are planned up to age 16 or 17 (see Schupp et al. 2008). This is the age when the "children" are interviewed as regular SOEP respondents. Given the panel character of the SOEP, such data is especially useful for longitudinal research, in particular for the analysis of preschool effects. With the SOEP, the relationship between preschool education, family indicators, and child outcomes can be analyzed. The SOEP data are available to the entire research community.

Another group of surveys consists of those with a research focus other than preschool education. One example of this type is the German Health Interview and Examination Survey for Children and Adolescents (KIGGS. *Studie zur Gesundheit von Kindern und Jugendlichen in Deutschland*) by the Robert Koch Institute.<sup>11</sup> It was designed as a nation-wide health survey for the age group from 0 to 17 years. Between 2003 and 2006, around 18,000

10 <http://www.diw.de/english/soep/26636.html>. See also Wagner et al. (2007).

11 <http://www.kiggs.de/>. See also Kurth et al. (2008).

children were enrolled. The data obtained include a large number of objective and subjective health measures. The parents were asked if the child visits a childcare program and the age of entry. Here, preschool education is valued as one of numerous other environmental determinants of child health. These data thus are useful for the analysis of the relationship between preschool education and child health. There will be a Public Use File of the KIGGS data available starting in 2009.

Apart from these representative datasets with a panel character, there exist other cross-sectional or regionally limited datasets that include information on preschool education. Most of these represent regional cross-sections with a special focus on children. Such a cross-sectional study is the DJI Children's Panel,<sup>12</sup> which collects detailed information about preschool education. In the study, more than 8,000 parents of children up to the age of six were interviewed. An example of a regionally restricted panel study is the German Longitudinal Study on the Genesis of Individual Competencies (LOGIK, *Longitudinalstudie zur Genese individueller Kompetenzen*), started in 1984 with 200 nearly four-year-old children in the region of Munich (see Weinert and Schneider 1999). The study, which was designed to analyze the development of various competences of children, contains information on preschool education as an environmental determinant of child development. Another study by Tietze et al. (1998) in three German states was designed to study the effects of preschool program quality. The study started with children in preschool and followed them up to primary school. Given the focus of this study, it is one of the few that has collected detailed information on the structural and process quality of preschool education. All these very different datasets were produced by a particular research group or institution with a particular research interest. In general, they are not available to the broader research community. Nevertheless, they can serve as models for the development of an adequate preschool data infrastructure.

From a European and international perspective, it is important to note that other countries realized the importance of a solid data infrastructure for research on preschool education – and on early education in general – long before Germany did. The Anglo-American research community was among the first to take up this issue in depth.<sup>13</sup> Thus, such major household panel studies as the US Panel Study of Income Dynamics (PSID) with its Child Development Supplement (CDS)<sup>14</sup> and the National Longitudinal Surveys of Youth (NLSY)<sup>15</sup> have special child-related supplements or questions. A number of these include research questions related to preschool education

---

12 <http://www.dji.de/cgi-bin/projekte/output.php?projekt=390>. See also Bien et al. (2006).

13 For a brief summary of Anglo-American longitudinal studies covering the early years, see BMBF (2008: chapter 2).

14 <http://psidonline.isr.umich.edu/CDS/>

15 <http://www.bls.gov/nls/>

that collect information on household and family context in particular. Since they cover some measures on the cognitive and non-cognitive development of children, they allow longitudinal studies on the effect of preschool education. Apart from major household panel studies, the NICHD study,<sup>16</sup> which began in 1991, is an example of a study offering an extremely rich database for the analysis of preschool education. The Cost, Quality and Child Outcomes Study in Child Care Centers (e.g., Helburn et al. 1995) is one of the few examples that combine detailed information on the quality of preschool programs and detailed information on its costs. There are a few studies that focus on particular model preschool programs. They allow a very detailed analysis of their effects, benefits, and costs (e.g., Perry Preschool Project, Schweinhardt et al. 2005). Great Britain is another country that acted on the perceived need for a data infrastructure on early education many years ago, producing various cohort studies starting at birth like the British Cohort Study,<sup>17</sup> the Millennium Cohort Study,<sup>18</sup> and the Avon Longitudinal Study of Pregnancy and Childhood (ALSPAC).<sup>19</sup>

For comparable European and international research, it would be beneficial for comparable survey instruments to be used in the collection of data on preschool education and child outcomes. The European Statistics on Income and Living Conditions (EU-SILC) is designed to provide harmonized instruments to this end. The EU-SILC data cover some preschool information, such as the number of hours in preschool.

### 3. Future developments

Given the well-recognized importance of preschool education, it is surprising that there are so few initiatives that exist in this field of educational research so far. In response to the general lack of empirically based educational research, the Federal Ministry of Education and Research (BMBF, *Bundesministerium für Bildung und Forschung*) has launched various initiatives to stimulate empirical research on education in Germany over the last several years. One of the few initiatives to focus on preschool (and school) education, however, is the BiKS Project.<sup>20</sup> BiKS stands for educational processes, competence development, and selection decisions at preschool and primary school age. It is based at the University of Bamberg and began the first round of data collection in 2005. The BiKS studies are based on formal

---

16 <http://secc.rti.org/>

17 <http://www.cls.ioe.ac.uk>

18 <http://www.millenniumcohort.org/>

19 <http://www.bristol.ac.uk/alspac>

20 <http://www.uni-bamberg.de/biks/>

childcare programs and schools in Bavaria and Hesse. The longitudinal study “BiKS 3-8” involves the observation of approximately 600 children, beginning at age three and continuing up to the second grade of elementary school.<sup>21</sup> This project will help to answer questions about the effects of preschooling, taking various socio-demographic and socio-economic variables of family background into account. In addition, the project will be concerned with the quality of preschool education.

Another project that reflects the efforts of the federal government to stimulate empirical educational research is the German National Educational Panel Study or NEPS.<sup>22</sup> The panel started in 2009. First information for children aged four and older will be collected. It is planned to collect information for younger children later. The sample will be drawn from institutions such as preschool programs, day care centers and schools. Once these data have been collected, they will provide the basis for innovative research, particularly on the effects of preschool programs. The NEPS offers a highly promising infrastructure, with a research team that includes social scientists from different disciplines such as sociology, education, psychology, and economics.

#### **4. Future developments: European and international challenges**

A European target has been set by the Barcelona European Council to provide childcare services for 90 percent of children between three years of age and the mandatory school age by 2010, and for 33 percent of children under three years of age (European Council 2002). In light of this, and even more recent initiatives from May 2005 that consider the improvement of childcare an important tool for meeting this target, the near-complete lack of internationally comparable data on preschool education is a significant problem.<sup>23</sup> The only exception to this is the information in EU-SILC but, as indicated above, this contains little preschool information and does not include quality data or child outcome data in a strict sense, such as variables on skills or socio-emotional behavior, which is precisely the type of data needed to study educational effects of preschool programs from a European

---

21 The longitudinal BiKS 8-12 follows approx. 2,000 children from third grade through sixth grade.

22 <http://www.uni-bamberg.de/neps>

23 The expansion of early childcare and education is also a goal on a broader international level. One of UNESCO's medium-term objectives (2002–2007) is the expansion and improvement of comprehensive early childcare and education, especially for the most vulnerable and disadvantaged children (UNESCO 2003).

perspective. Apart from this, it is not clear what the different countries mean by the term “childcare” and thus whether this information can be used for research on preschool education.

In light of all these findings, it is clear that *there is need in most countries for a systematic procedure* to collect and provide consistent and comparable information on preschool education programs. Currently, the ministries responsible for young children use different indicators and diverse methods in collecting data on the preschool education of young children. Thus the definition of the population group considered to be in pre-primary education is often arbitrary. Moreover, program criteria are sometimes confusing. What is clear is that countries use different proxy measures to determine whether a program should be classified as educational or not. Variation in these proxy measures undermines comparability. Moreover, the weekly and annual durations of preschool education sessions are rarely taken into account. Thus it is obvious that the first future challenge to be faced is to provide comparable data on preschool education with respect to basic structural characteristics. Once these data exist, a further challenge will be to provide data on program quality, the costs and potential outcomes of preschool education programs in the European countries, or at an even broader level among the various OECD countries (on the need for comparable data, see OECD 2006, esp. chapter 8).

## **5. Conclusions and recommendations**

Given the importance of improved preschool programs in Germany and taking into account the various initiatives underway on different levels, the following demands for a data infrastructure in this area can be summarized: (1) Current and detailed data on the development of preschool programs is needed (detailed with respect to age groups, hours of care provided, provider type, etc.). This is also needed to ensure that the current political initiatives are effective. This data should be available on a small regional level since there are great regional disparities that needed to be analyzed. (2) Data on attendance in these preschool programs are needed that cover the socio-demographic and socio-economic backgrounds of the children and their families. This will make it possible to address important questions such as whether disadvantaged children attend preschool. (3) Given the importance of preschool programs, data on the quality of these programs should be made available. These data should not focus on structural quality indicators alone (e.g., group size, child-staff ratios). As research has shown, other quality dimensions such as process quality (the interaction between child and teacher) are even more important for child development (see, for instance,

Roßbach 2005). (4) These data should ideally be linked to cost information. General information on the expenses of such programs is important, but not necessarily sufficient. What is ideally needed is data on the detailed costs of particular programs in relation to the level of educational quality. (5) Especially from a longitudinal point of view, it is important to have child outcome data<sup>24</sup> that can be linked to preschool information. This child outcome data should include cognitive measures as well as socio-emotional outcome measures. Only this linkage allows the effects of preschool programs to be analyzed in the short, medium, and long term. Given the fact that most of the empirical research on the effects of preschool programs has been conducted in Anglo-American settings, this type of analysis is widely missing in Germany, where the necessary microdata to support it is virtually non-existent. Up to now, there have only been few empirical studies on preschool effects in Germany, partly based on regionally restricted samples. Moreover, only a few of the existing studies were able to control for the quality of the preschool programs and even less for their costs (for a short summary of these studies based on representative microdata see Spieß 2008a).

Given the recent developments in German data infrastructure, the potential for preschool education research will certainly improve. Nevertheless, a number of gaps remain:

- (1) First of all, groups of children under three or four years of age have not received adequate attention to date. This partly reflects the longstanding idea that preschool education begins with entrance to the traditional German *Kindergarten* at age three. Yet developmental psychologists and brain researchers have shown that education starts much earlier. Moreover, there is a new political commitment to increasing preschool attendance of children below the age of three. It is therefore recommended that the data infrastructure be improved particularly around this early stage of child development.<sup>25</sup>
- (2) Given the importance of preschool for disadvantaged children – such as children from households with low socio-economic status – datasets covering this group should have an adequate sample size. An over-sampling of this group in existing surveys might be one option; special surveys of these groups of children might be another. Nevertheless, future study designs should integrate the setup of a control group as well.

---

24 For a summary of relevant child outcome indicators with respect to various competencies in early childhood, see, for instance, BMBF (2008).

25 The SOEP data are almost the only publicly available dataset covering this early age. More data is also needed on family day care, which will probably play an increasingly important role in the future. Therefore, efforts to improve the data infrastructure for the younger age group should be linked with efforts to improve the data situation on family day care.

- (3) Third, there are no representative data on the quality of preschool education programs if researchers do not want to rely solely on structural quality indicators (such as group size, staff-to-child ratio, or education of the staff). This is striking, considering that other quality dimensions, such as process quality, have been shown to be of much greater importance than structural indicators in explaining the variance in child outcomes. However, the lack of representative data on this aspect might be related to difficulties inherent in measuring process quality. There are a few instruments for quantifying process quality (for a summary, see McCabe and Ackermann 2007), but they are costly and time intensive. Thus, on the one hand, greater effort should be put into the development of more efficient instruments for measuring process quality with respect to cost and time. On the other hand, more effort should be put into the application of existing instruments to a broader set of preschool programs.
- (4) Fourth, given the significant impact of preschool quality on child development, measures of quality should be available in datasets that also include child outcome measures. Given the lack of efficient measures of process quality in the short-run, structural quality indicators can be used in a first stage. Official data such as the Children and Youth Services Statistic cover some indicators on the structural quality of preschool programs, while surveys such as the SOEP cover child outcome measures. At the moment, however, it is not possible to link this information. Therefore, serious investigation needs to be undertaken to determine how survey data can be linked with official data, and it is obvious that the preschool institution a child attends is the key indicator. The efforts of labor market researchers to link official data with survey data in their field of research might provide a useful model. Another option would be to enrich datasets like the SOEP with quality data on preschool by collecting additional quality data on preschool for the explicit purpose of linkage. This is possible once the respondent agrees to allow the institution that his or her child attends to be identified.<sup>26</sup> It is obvious that data security will play a major role in such an undertaking. But such efforts are of particular interest in the long run since they enable long-term analysis.
- (5) Fifth, the crucial interaction between the family or home environment and preschool education can be analyzed using current and future data, but only to a very limited extent. This is particularly true with respect to the quality of these two educational settings, although it is known from international research that these quality aspects are of extraordinary

---

26 The SOEP has conducted a pretest asking respondents for information on the preschool or day care center that their child attends (see Schupp et al. 2008: 72).

importance. It is therefore recommended that greater effort be expended to provide data on the quality of the family/home environment<sup>27</sup> and the quality of preschool education simultaneously. Such data would be of particular interest for multilevel analysis.

- (6) Apart from preschool quality and child outcome aspects, there is a dearth of information on the real costs of preschool programs. There are almost no options for combining detailed cost information with preschool quality and child outcome information. This is remarkable since pure information on cost is not very useful in itself for educational research, and the relationship between preschool quality and child outcomes and a given input (here, particular costs) is unknown. Better data on costs, preschool quality, and child outcomes are necessary for solid cost-benefits analyses of the German preschool system. A cost-benefit analysis based on such data would complete the set of first attempts for cost-benefit calculations in Germany (see Spieß 2008b).
- (7) Nevertheless, it is clear that compiling nationally representative datasets that cover detailed cost information, quality information, and detailed information on various aspects of the cognitive and non-cognitive development of children is an extremely costly and time-consuming endeavor. To save resources, it might be possible to add the missing information to a preexisting dataset at a less detailed level. Alternately, one could focus on particular preschool programs or a particular group of children. In this case, we would recommend more intervention studies in the strict sense. These studies could have different foci, but should all aim at collecting more detailed information on preschool quality, child outcomes, and costs. They would also allow us to learn more about the effects of specific preschool programs with a special educational program.
- (8) From an international, and in particular European perspective, there is a clear need for more, and more comparable data on preschool education in the various countries. As pointed out in section 4, the primary aim is to collect and provide more detailed information on preschool, with information on structural indicators. Such *datasets need to cover newborns to six-year-olds*, and include all forms of provision, regardless of administrative responsibility, funding source, or setting. What is needed is a collection of data over time. From a longer-term perspective, comparable data are needed, which offer more information on the quality, costs, and outcomes of such preschool education programs.

---

27 One example of a scale that could be used to measure the quality of a child's home environment, see the HOME scale as used in the NLSY (Bradely et al. 2001).



Apart from this, there is a special need for collecting and identifying preschool information for children from disadvantaged families.

- (9) All these recommendations reveal obvious links to the current state of the data infrastructure in the field of families in general (see Huinink in this publication) and in the fields of abilities and competencies (see Stern, Trautwein and Schoon in this publication). Thus it might be of added value to develop the preschool data infrastructure in collaboration with the data infrastructure of the other fields mentioned. Although the foci might differ, the overlap between these various fields should be kept in mind and efforts should be made to foster exchange between the different interests, agencies, and organizations involved.

## References:

- Alt, C. (2005): Data, Design and Constructs – The First Wave of the Children's Panel. In: Klöckner, C. and Paetzel, U. (Eds.): *Kindheitsforschung und kommunale Praxis. Praxisnahe Erkenntnisse aus der aktuellen Kindheitsforschung*. Wiesbaden.
- Bien, W./Rauschenbach, T. and Riedel, B. (Eds.) (2006): *Wer betreut Deutschlands Kinder?* Weinheim/Basel.
- Bradley, R.H./Corwyn, R.F./McAdoo, H.P. and Coll, C.G. (2001): The Home Environments of Children in the United States Part I: Variations by Age, Ethnicity, and Poverty Status. *Child Development* 72, 1844-1867.
- BMBF – Bundesministerium für Bildung und Forschung (Ed.) (2008): *Kindliche Kompetenzen im Elementarbereich: Förderbarkeit, Bedeutung und Messung. Reihe Bildungsforschung des BMBF, Volume 24*.
- European Council (2002): Presidency conclusions, Barcelona European Council, 15-16 March 2002, available at: [http://ue.eu.int/ueDocs/cms\\_Data/docs/pressData/en/ec/71025.pdf](http://ue.eu.int/ueDocs/cms_Data/docs/pressData/en/ec/71025.pdf). [Last visited 02/03/2010].
- Helburn, S. (1995): Cost, Quality, and Child Outcomes in Child Care Centers: Key Findings and Recommendations. *Young Children* 50, 40-44.
- Roßbach, H.-G. (2005): Effekte qualitativ guter Betreuung, Bildung und Erziehung im frühen Kindesalter auf Kinder und ihre Familien. In: Sachverständigenkommission Zwölfter Kinder- und Jugendbericht (Ed.): *Bildung, Betreuung und Erziehung von Kindern unter sechs Jahren*. München.
- Karoly, L.A./Kilburn, R.M. and Cannon, J.S. (2006): *Early Childhood Interventions: Proven Results, Future Promise, RAND Cooperation*. Santa Monica.
- Kolvenbach, F.-J. and Taubmann, D. (2006): Neue Statistiken zur Kindertagesbetreuung. *Wirtschaft und Statistik* 2, 166-171.
- Konsortium Bildungsberichterstattung (Ed.) (2006): *Bildung in Deutschland*. Gütersloh.
- Kurth, B.-M. et al. (2008): The challenge of comprehensively mapping children's health in a nation-wide health survey: Design of the German KiGGS-Study. *BMC Public Health* 8, 196.
- Lohmann, H./Spieß, C.K./Groh-Samberg, O. and Schupp, J. (2008): Analysepotenziale des Sozio-oekonomischen Panels (SOEP) für die empirische Bildungsforschung. SOEP Paper No. 110.
- McCabe, L.A. and Ackerman, D.J. (2007): *Child Care Center Quality: Measurement Issues and Links to Child Developmental Outcomes*. DIW Data Documentation No. 25.
- OECD (2006): *Starting Strong II*. Paris.
- OECD (2007): *Babies and Bosses. A Synthesis of Findings for OECD Countries*. Paris.
- Roßbach, H.-G. (2005): Effekte qualitativ guter Betreuung, Bildung und Erziehung im frühen Kindesalter auf Kinder und ihre Familien. In: Sachverständigenkommission Zwölfter Kinder- und Jugendbericht (Ed.): *Bildung, Betreuung und Erziehung von Kindern unter sechs Jahren*. Bd. 1. München.

- Schilling, M. (2002): Die amtliche Kinder- und Jugendhilfestatistik. Dissertation am Fachbereich Erziehungswissenschaft und Soziologie der Universität Dortmund, November 2002.
- Schupp, J./Spieß, C.K. and Wagner, G.G. (2008): Die verhaltenswissenschaftliche Weiterentwicklung des Erhebungsprogramms des SOEP. Vierteljahrshefte zur Wirtschaftsforschung 77, 63-76.
- Schweinhardt, L.J./Montie, J./Xiang, Z./Barnett, W.S./Belfield, C.R. and Nores, M. (2005): Lifetime Effects: The High/Scope Perry Preschool Study Through Age 40. Ypsilanti.
- Spieß, C.K. (2008a): Early Childhood Education and Care in Germany: The Status Quo and Reform Proposals. Zeitschrift für Betriebswirtschaftslehre 67, 1-20.
- Spieß, C.K. (2008b): Volkswirtschaftliche Bedeutung der Kinderbetreuung: Wie ist diese zu bewerten und was können wir dabei aus dem Ausland lernen? In: von der Leyen, U. (Ed.) (2008): Voneinander lernen – miteinander handeln: Aufgaben und Perspektiven der Europäischen Allianz für Familien. Berlin.
- Tietze, W. (Ed.) (1998): Wie gut sind unsere Kindergärten? Eine Untersuchung zur pädagogischen Qualität in deutschen Kindergärten. Weinheim.
- Wagner, G.G./Frick, J.R. and Schupp, J. (2007): The German Socio-Economic Panel Study (SOEP): Scope, Evolution and Enhancements. Schmollers Jahrbuch 127, 139-170.
- Weinert, F.E. and Schneider, W. (Eds.) (1999): Individual development from 3 to 12: Findings from the Munich Longitudinal Study. New York.
- UNESCO (2003): Medium-term strategy 2002-2007, available at: <http://unesdoc.unesco.org/images/0012/001254/125434e.pdf>. [Last visited 02/03/2010].
- Unicef Innocenti Research Centre (UNICEF) (2008): The child care transition, Report Card 8. Florence.



### 3.3 Data in the Domain of Secondary School Education

*Petra Stanat and Hans Döbert*

Contact:

Petra Stanat  
Humboldt-Universität zu Berlin  
Institut zur Qualitätsentwicklung im Bildungswesen  
Unter den Linden 6  
10099 Berlin  
Germany  
e-mail: [petra.stanat\[at\]iqb.hu-berlin.de](mailto:petra.stanat[at]iqb.hu-berlin.de)

## Abstract

Research on school education is exceptionally active at present. This heightened level of activity is partly due to the realization that, compared to other countries, Germany knows very little about its school system. Before the results from the first cycle of the Programme for International Student Assessment (PISA) were published at the end of 2001, for example, even the proportion of immigrant students attending German schools was largely unknown (Baumert et al. 2001). Although the situation has changed tremendously over the last 10 years, many questions remain open. One of the major research gaps pertains to how student competencies and other aspects of educational success develop over time and across different stages of the education system. Similarly, information on the factors that shape these developments is lacking. This is particularly the case for process factors within schools, classrooms, and families that affect student learning. Moreover, although considerable progress has been made in capturing cognitive competencies and skills, little is known about how they unfold over time. Also, the role that “soft-skills,” such as social competencies, play as determinants and outcomes of educational processes remains largely unclear. To provide a basis for exploring these and other issues, it is necessary to make existing datasets available to researchers and to generate additional datasets with improved research designs and instrumentation.

## 1. Currently available datasets

Three types of datasets are currently available in the domain of school education in Germany: official and non-official school statistics (primarily the school statistics and the Microcensus), survey data (e.g., SOEP,<sup>1</sup> ALLBUS,<sup>2</sup> the HIS<sup>3</sup> survey of students eligible for university studies), and data from large-scale assessments of student performance (e.g., PISA, TIMSS,<sup>4</sup> PIRLS,<sup>5</sup> *Länder* assessments).<sup>6</sup> From the perspectives of educational monitoring and reporting, the usefulness of these datasets depends, among other things, on the extent to which they include information on (1) student educational participation (including the type of school attended within the

---

1 German Socio-Economic Panel (SOEP, *Sozio-oekonomisches Panel*).

2 German General Social Survey (ALLBUS, *Allgemeine Bevölkerungsumfrage*).

3 Higher Education Information System (HIS, *Hochschul-Informationssystem*).

4 Trends in International Mathematics and Science Study.

5 Progress in International Reading Literacy Study.

6 Important additional data sources are research projects focusing on specific aspects of the school system, such as video studies on classroom instruction (e.g., Seidel et al. 2007), studies on transitions from elementary to secondary school (e.g., Ditton 2007), or studies on whole day schooling within the Study on the Development of Whole-day Schools (StEG, *Studie zur Entwicklung von Ganztagschulen*) program (Holtappels et al. 2008). It is beyond the scope of this report, however, to cover these studies as well.

tracked system and grade retention), (2) learning outcomes in core domains, (3) family background, and (4) student development over time (preferably based on longitudinal data rather than on retrospective information). In terms of these features, each of the listed dataset has differential strengths and weaknesses.

- The main advantage of *school statistics* is that they are based on data collected for the entire school population. They include information on educational participation and grade retention. However, the data are cross-sectional and are provided at the aggregate level only. It is consequently neither possible to perform individual-level analyses nor longitudinal analyses based on school statistics. In terms of student background, the data include only rudimentary information, such as gender and nationality. Moreover, indicators of learning outcomes are not available. An attempt is currently being made to reform the data collection process for school statistics. This reform, which will be described in more detail below, would be in line with the approach taken in many other countries where school statistics are already based on individual-level data and include a student identification number.
- Although the *Microcensus* provides data on individuals within households, its usefulness for school-related analyses is also quite limited. In the past, the Microcensus asked respondents only whether or not household members attended a school and, if so, which grade they were in. Information on the type of school attended was not collected. This changed in the 2008 survey, which now includes questions on the types of school the household members visit. At the same time, however, questions on attendance of preschool institutions, such as kindergarten, have been eliminated from the survey. Moreover, although the Microcensus now collects more detailed information on family immigration background, some of the most important questions (e.g., country of birth) will only be asked every four years. Like school statistics, Microcensus data are cross-sectional and do not include any indicators of learning outcomes. Similar household surveys are also conducted in other countries, such as the Labour Force Survey (LFS) in the UK or the Swiss Labour Force Survey (SAKE, *Arbeitskräfteerhebung*).
- Among all the surveys carried out in Germany, the *Socio-Economic Panel (SOEP)* is the most frequently used dataset for individual-level analyses related to schooling. It provides information on school participation and on a large number of background factors over time. Since 2006, the SOEP has also collected data on various aspects of cognitive functioning (Wagner et al. 2007). Adolescents are tested with a measure of verbal, numerical, and figural intelligence, and adults complete short tests on processing speed and word fluency. This addition is useful for



many analyses, yet its relevance for questions related to outcomes of schooling is limited. Although schools certainly affect students' intelligence development, their effects on such domain-general cognitive dispositions are considerably smaller than on subject-specific competencies, such as mathematics or foreign languages (Baumert et al. 2007). Again, similar longitudinal household surveys are carried out in other countries, such as the Longitudinal Household Study in the UK (UKLHS) or Labourmarket Monitoring (LAMO) in Austria.

- *Large-scale assessments of student performance* have gained in importance considerably over the past ten years. Due to factors such as a generally negative attitude towards testing that was shared by many members of relevant stakeholder groups in Germany, as well as the potential conflict associated with comparisons of student performance across the federal states, or *Länder*, Germany has, for a long time, refrained from measuring the output of schooling. In the aftermath of the first cycle of the OECD's Programme for International Student Assessment (PISA), whose results were published in 2001 (Baumert et al. 2001; OECD 2001), however, this changed dramatically. The disappointing results that PISA revealed for Germany has spurred a national paradigm shift, from an almost exclusive input-orientation in the school system to a considerably stronger focus on its output – a shift that many other countries, such as the Netherlands, Sweden, or the United States, have undertaken long ago. The Standing Conference of the Ministers of Education and Cultural Affairs of the *Länder* in the Federal Republic of Germany (KMK, *Kultusministerkonferenz*) decided to participate regularly in the international large-scale assessment studies PISA, TIMSS elementary school, and PIRLS. In addition, the KMK set up the Institute for Educational Progress (IQB, *Institut zur Qualitätsentwicklung im Bildungswesen*) whose task it is to coordinate the specification of national standards for learning outcomes in various school subjects and to develop test instruments that can be used to evaluate the extent to which the standards are met (Köller 2008). The IQB will administer these tests in representative samples in order to provide information on student performance levels in the 16 *Länder*.

In addition to these national activities, several of the German states have carried out their own assessment studies (e.g., LAU,<sup>7</sup> KESS,<sup>8</sup> MARKUS,<sup>9</sup>

7 Aspects of Learning Prerequisites and Learning Development (LAU, *Aspekte der Lernausgangslage und der Lernentwicklung*).

8 Competencies and Attributes of Students (KESS, *Kompetenzen und Einstellungen von Schülerinnen und Schülern*).

9 Rhineland-Palatinate Comprehensive Assessment in Mathematics: Competencies, Characteristics of Instruction, School Context (MARKUS, *Mathematik-Gesamterhebung Rheinland-Pfalz: Kompetenzen, Unterrichtsmerkmale, Schulkontext*).

ELEMENT,<sup>10</sup> QuaSUM<sup>11</sup>), and all of them are currently conducting “comparison tests” or *Vergleichsarbeiten* on a regular basis in selected grades (typically grades 3 and 8). The main purpose of the *Vergleichsarbeiten* is to provide schools, teachers, parents, and students with feedback on their relative learning results (Hosenfeld, 2008). Although the tests are developed centrally, they are administered and scored by the students’ teachers, so the quality of the data is unclear. Moreover, whereas the international, national, and *Länder*-specific assessment studies typically measure a wide range of background variables, this type of information is either absent from or highly limited in the *Vergleichsarbeiten* datasets.

The international and national student assessment studies present an important data source for analyses on secondary schools. They provide highly reliable information on educational participation as well as learning outcomes. With the exception of a national extension to PISA 2003 (Prenzel et al. 2006) and the nationwide study on language competencies DESI<sup>12</sup> (Klieme et al. 2008), both of which included a longitudinal addition with two measurement points (PISA: from the end of grade 8 to the end of grade 9; DESI: from the beginning to the end of grade 9), the national assessments typically have cross-sectional designs. Longitudinal studies in selected regions of Germany include: BIJU<sup>13</sup> (grades 7–12 and transition to work), BiKS<sup>14</sup> (ages 3–8 and 8–12), DESI (beginning and end of grade 9), ELEMENT (grades 4–10), LAU<sup>15</sup>/ULME<sup>16</sup> (grades 5–13, vocational training), SCHOLASTIK<sup>17</sup> (grades 1–4), and TOSCA<sup>18</sup> (grade 10–vocational training or university).

---

10 Study on Competencies in Reading and Mathematics (ELEMENT, *Erhebung zum Lese- und Mathematikverständnis*).

11 Study on School Quality in Mathematics Instruction (QuaSUM, *Qualitätsuntersuchung an Schulen zum Unterricht in Mathematik*).

12 Assessment of Student Achievements in German and English as a Foreign Language (DESI, *Deutsch-Englisch-Schülerleistungen-International*).

13 Learning Processes, Educational Careers, and Psychosocial Development in Adolescence and Young Adulthood (BIJU, *Bildungsverläufe und psychosoziale Entwicklung im Jugend- und jungen Erwachsenenalter*).

14 Educational Processes, Competence Development, and Selection Decisions in Pre- and Primary School Age (BiKS, *Bildungsprozesse, Kompetenzentwicklungen und Selektionsentscheidungen im Vor- und Grundschulalter*).

15 Aspects of Learning Prerequisites and Learning Development (LAU, *Aspekte der Lernaussgangslage und der Lernentwicklung*).

16 Study on Achievement, Motivation, and Attitudes of Students at the Beginning of Vocational Education (ULME, *Untersuchungen der Leistungen, Motivation und Einstellungen zu Beginn der beruflichen Ausbildung*).

17 Organized Learning Opportunities at School and the Socialization of Talents, Interests, and Competencies (SCHOLASTIK, *Schulorganisierte Lernangebote und die Sozialisation von Talenten, Interessen und Kompetenzen*).

18 Transformations of the Secondary School System and Academic Careers (TOSCA, *Transformation des Sekundarschulsystems und akademische Karrieren*).

In general, the types of datasets that are currently available in Germany typically also exist in other countries. Germany lags behind some international developments, yet the process of catching up with these developments is currently well under way.

## 2. New developments

### 2.1 *Change of school statistics to individual-level data*

In 2003, the KMK decided to change the school statistics to individual-level data and defined a core set of variables to be included in the dataset (“Kern-datensatz für schulstatistische Individualdaten der Länder”). The core dataset, which the individual *Länder* may extend as they see fit, encompasses the following variables:

- *Organizational characteristics of the school* (e.g., location, school type, legal status, number of delayed enrollments in first grade)
- *Individual background data of students* (e.g., sex, month and year of birth, nationality, year of immigration to Germany, language spoken in the family, country of birth, grade level, year of enrollment in first grade, school and grade attended the year before, type of grade repetition, focus of special support measures, attendance of all-day schooling programs and after-school care, place of residence)
- *Individual background data of school-leavers and graduates* (similar to data for students remaining in the school as listed above)
- *Individual background data of teachers and data on teacher fluctuation* (e.g., sex, month and year of birth, nationality, type of teacher training, teaching qualification by subjects, gross teaching load, type and hours of reduced or additional teaching load, number of lessons taught at the school, functions within the school’s administration, beginning and end of employment at the school)
- *Data on classes/courses at the school*
- *Organizational data in terms of instructional units* (e.g., grade level, educational track, subject, lessons per week, course type)

The KMK had aimed to convert school statistics to individual-level data with the agreed-upon core set of variables by 2007. Since the process has not been finalized across the board, however, the time frame has been extended to 2008–2009. Some of the *Länder* already possess individual datasets with student identification numbers while others have not yet begun to change

their statistics. Moreover, the *Länder* typically concentrate on public schools, and it is not clear whether the collection of individual school data will be extended to private schools in the future.

The conversion of school statistics to individual-level data is in line with international standards. Informal interviews with representatives from several Northern European countries (including Austria and Switzerland) in the technical group of the OECD's Educational Indicators Project indicate that individual-level data are the rule rather than the exception (Hetmeier and Leidel 2007). Among the eleven countries included in the interviews, five (Austria, Denmark, Israel, the Netherlands, Switzerland) collect individual-level data for students in elementary schools and all countries except Poland and Spain collect such data for students in secondary schools. Moreover, the datasets in these countries typically include identification numbers, making it possible to capture educational careers over time. Austria, the Netherlands, and – starting in 2011/2012 – Switzerland use the social security number for this purpose; Denmark, Finland, Norway, Poland, and Sweden use a general person registration number.

In Germany, the data protection commissioners of some *Länder* resist the introduction of pupil identification numbers as well as the establishment of a centralized national data pool. In an attempt to find a compromise solution for this issue, a hashcode is considered that would be derived from data that typically remains unchanged throughout a student's school career (e.g., date of birth, sex, name), thus making it possible to capture educational pathways longitudinally. However, it is highly uncertain whether such a solution will, in fact, be accepted and implemented by all *Länder*. It is also unclear whether individual-level data will be integrated into a national data pool and whether it will cover the entire school system, including private schools and vocational schools.

## 2.2 *The National Educational Panel Study*

Carried out by a consortium of researchers from various disciplines concerned with education, its outcomes and returns (education, psychology, sociology, and economics), the National Educational Panel Study (NEPS) promises to provide a dataset that meets the criteria outlined above (Blossfeld et al. 2008). It will include comprehensive information on the participants' educational careers, competencies as well as family background, and it will collect this information longitudinally. The panel is designed as a multi-cohort sequence study with eight stages that cover important transitions within the educational system. Secondary schooling will be analyzed in stage 3 (from elementary school to lower secondary school), in stage 4 (from lower secondary to upper secondary school), and in stage 5 (from lower secondary school to higher education, occupational training, or the labor market). The

assessments will focus on five themes that are of particular relevance within research on education. These are (1) competence development across the life-span, including German-language competencies and reading literacy, mathematical literacy, scientific literacy, foreign-language competencies, and social competencies; (2) educational processes in life course specific learning environments; (3) social inequality and educational decisions in the life course; (4) education acquisitions of immigrants and their descendants across the life span; and (5) education returns. The first assessments within the NEPS will start in 2009–2010 so that the results can be expected to become available in 2011–12.

The implementation of the NEPS is also in line with international developments. Several countries already have similar longitudinal studies in place, such as the 1970 Birth Cohort, the Millennium Cohort, and the Youth Cohort Studies in the UK; the National Longitudinal Survey of Children and Adolescents (NLSCY) in Canada; or the Survey of Children and Youth on “Competence and Context” (COCON) in Switzerland. However, the scope of these studies tends to be more limited than that of the German NEPS, both in terms of the age range included in the assessments and in terms of the covered research foci. A particularly innovative and challenging feature of the NEPS is that it aims at assessing central domain-specific competencies across the lifespan. In addition, while it would be possible to study educational careers based on adequate individual-level school statistics, NEPS will allow for more detailed and in-depth analyses of these careers within contexts. Thus, the NEPS is a highly ambitious program which goes beyond what is currently available internationally. If successful, the study will yield an important database for analyses of educational processes and outcomes.

### **3. Data access**

Data from the Microcensus and such surveys as the SOEP or ALLBUS are typically available as Scientific Use Files. Within the limits of data protection regulations, researchers are free to use the data for scientific analyses. Similarly, the NEPS data will be made accessible to the scientific community as soon as the necessary data cleaning and scaling procedures have been completed. While the international dataset from such studies as TIMSS and PISA can be easily downloaded from the Internet, access to national data from assessments of student performance has traditionally been more restricted. The Federal Ministry of Education and Research (BMBF, *Bundesministerium für Bildung und Forschung*) is funding a Research Data Center at the IQB that is designed to make data from student assessment studies

available to researchers and to provide training and support in using these data (IQB 2009). Researchers who want to work with a specific dataset have to submit a proposal, describing their research questions and the analyses they plan to carry out. Whether or not access is granted depends on the following criteria (IQB 2008):

- (1) The data will be used for scientific purposes, not commercially.
- (2) Individual data protection is secured.
- (3) The planned analyses are in line with contractual agreements made with the owner of the data (such as the KMK). Comparisons between the German *Länder*, which have not yet been conducted, need to be approved by the KMK.
- (4) The planned analyses do not threaten theses (such as dissertations) or publications that are currently being written. The topics of these projects have to be specified at the time the researchers who have collected the data submit them to the Research Data Center. These topics are blocked for analyses with the respective dataset for at least nine months. This time lag may be extended to three years at the most.
- (5) No additional issues are in conflict with the intended use of the data.

If these conditions are met, the Research Data Center at the IQB will provide the applicant with a Scientific Use File. Researchers who want to use sensitive data (e.g., the *Länder* codes) have to perform their analyses at the IQB or via remote computing.

Unlike the initial regulations of the Research Data Center at the IQB, which included an evaluation of proposals in terms of their theoretical and methodological soundness, the current procedure largely complies with the standards for Research Data Centers defined by the RatSWD. Nevertheless, a few open questions remain, such as who decides whether or not a proposal “threatens” theses or publications on the blocked topics and what kind of “additional issues” may be “in conflict with the intended use of the data.” The transparency of the decision-making process should be increased by publishing a list of specified research questions that would be rejected on the grounds of the third or fourth criteria. This would prevent researchers from investing time and effort on writing proposals that are bound to fail, and it would further increase the perceived fairness of the application process.

#### 4. Current challenges and future development

Although the database for analyses on secondary schooling has improved considerably, a number of challenges remain. Most of these issues pose significant challenges not only in Germany but also at the international level, such as the problems associated with modeling input-process-output associations in education.

Variables in educational research can be categorized in terms of whether they pertain to the input, to the process, or to the output of education (Konsortium Bildungsberichterstattung, 2006). While more work is needed on the systematization and operationalization of variables in all three categories, the process dimension is the most challenging. From the perspectives of both educational research and educational reporting, it would be highly desirable to generate data and indicators that capture the process character of education. This pertains to characteristics of *interactions* in educational settings that are related to actual learning processes, such as approaches to structuring the teaching and learning process or the use of instruction time by teachers and learners, as well as to aspects directly relevant to the governance of educational *institutions*, such as the implementation of curricular requirements or measures of quality development and quality control. The central question is whether it is possible to develop indicators for basic dimensions of processes that can be measured across different educational institutions and across the sixteen *Länder* in reliable and valid ways. Since such measures are necessary to study the black box between the input and the output of education, it would certainly be worthwhile to invest into their development.

Within the output dimension, further work is needed on the assessment of so-called “soft skills,” such as the various facets of social competence (e.g., the ability to communicate and cooperate with others). There is unanimous agreement that these skills present important determinants of learning processes as well as outcomes of schooling, and it has repeatedly been argued that they should be taken into account in educational monitoring. Thus far, however, no reliable and valid measure has been developed that could be included in such studies as PISA (Kanning 2003; Kunter and Stanat 2002). This is partly due to the fact that the appropriateness of social behavior is highly dependent on situational requirements; that is, a specific behavior may be quite competent in one situation yet largely counterproductive in another. Because it is typically not feasible to employ systematic observations in large-scale surveys or assessment studies, it would be necessary to develop more indirect measures. One promising approach might be to work with computerized scenarios that elicit student responses to simulations of various social situations. Yet, again, it would be a challenge to come up with scoring systems for the appropriateness of these reactions.

Even aspects of the output dimension for which well-established measures exist (e.g., reading, mathematics, and science), however, require further development for long-term longitudinal analyses. Most of the existing instruments were specifically designed for groups of students within a restricted age range, such as students in 9<sup>th</sup> grade or 15-year-olds. A major challenge the NEPS will have to tackle is to model these competencies and link the measures across the life span. If successful, however, the study will provide important insights into the ways in which competencies in the domains of language, mathematics, and science unfold and interrelate at different developmental stages.

The question of how to model change is a challenge for educational research that has only been partly resolved. It is relevant not only at the level of individual development, but also for the development of individual schools and school systems. In Germany, there is a dearth of research on the extent to which schools change over time and on factors determining this change (Klieme and Steiner 2008). This is mainly due to the lack of longitudinal data at the school-level. Although the comparative tests (*Vergleichsarbeiten*) discussed above could be used to perform such analyses, their results should be validated through studies employing more controlled data collection and scoring procedures. Several analyses of this type will be necessary to derive reliable estimates of changes occurring at the institutional level, the stability of these changes, and the effect sizes associated with potential determinants of developmental trajectories.

Even more complex is the attempt to capture and explain change at the level of school systems. The trend design of PISA aims at providing information on the extent to which school outcomes in the participating countries improve or do not improve over time, yet the interpretation of the findings has been controversial (Carstensen et al. 2008). This complexity is partly due to the changing focus of the PISA assessments, such that in each project cycle one of the three assessment domains (reading, mathematics, and science) is measured more comprehensively than the other, and to the multi-matrix design employed in the study. In PISA 2000 the focus was on reading, in PISA 2003 on mathematics, and in PISA 2006 on science. As a consequence, the overlap of items across the cycles has, so far, been limited. In addition, the composition of the test booklets varied in the different cycles, making it difficult to tease apart possible changes in item difficulty and changes in performance levels.

Still more complicated than estimating changes at the levels of schools and school systems is the attempt to explain the observed developments with multilevel analyses. Multilevel modeling presupposes that the variables included in the model are comparable across the units of analyses. This is particularly questionable for analyses of data from international assessment studies, as specific features of individual countries, such as approaches to



ability grouping or types of curricula within the tracks, will almost inevitably be neglected (see Stanat and Lüdtke 2008 for a discussion of multilevel issues in international assessment studies). It is typically impossible to take such complex between-country differences into account in multilevel analyses. Therefore, quantitative multilevel analyses need to be complemented with qualitative data in order to generate more in-depth information on single cases (e.g., Döbert and Sroka 2004). One approach, for example, would be to submit countries deviating substantially from their predicted value in a multilevel analysis (Bowers and Drake 2005) to an intensive ideographical analyses. In addition to structural features of the school system, these analyses should take into account the country's historical complexity as well as cultural factors that are likely to affect teaching and learning processes (Stanat and Lüdtke 2007).

## 5. Recommendations

Based on the current situation outlined above, the following recommendations result:

### School statistics

- School statistics should be changed to individual-level data in all *Länder*.
- To allow for longitudinal analyses of educational careers, school statistics should include a student identification code that remains the same across core educational stages. In the short run, the feasibility of different approaches to deriving such a code should be evaluated.
- Core characteristics, such as student socio-economic and immigrant backgrounds, should be represented by the same indicators in all *Länder*.
- More generally, the measures of core characteristics should be harmonized in the available statistical datasets (e.g., school statistics, youth welfare statistics, vocational training statistics).
- Data should also be collected for private schools.

### Large-scale assessments (with *Länder* comparisons)

- Participation in international large-scale assessment studies, especially PIRLS, TIMSS, and PISA should be continued.

- In order to use the potential of these studies more fully, it should remain possible to add national options to the international designs.
- Assessments of competencies should be extended to the beginning of elementary school (including language skills) and to the transition into the labor market.
- In addition to the large-scale assessment studies, the German Time Use Study (*Zeitbudgetstudie*) and the Volunteer Survey (*Freiwilligensurvey*) should be continued as well. The Time Use Study is the only source of reliable data on the time people invest in education, and the Volunteer Survey allows for analyses of relationships between background factors and non-formal as well informal learning.
- Data from large-scale assessments and similar studies should be made available to the scientific community as soon as possible after they have been collected, cleaned, and scaled.
- The procedure of granting individual researchers access to data from large-scale assessments and similar studies should be completely transparent, including openness about the limitations of access in terms of the content of the proposed analyses. Such content-related limitations should be avoided.

#### Conceptual work and instrument development

- Theoretical models specifying the structure of competencies need to be refined as well as tested in various domains.
- In some domains, such as social or vocational competencies, the development of conceptual models is still in its very early stages and should be intensified.
- Substantive conceptual work is also needed with regard to the theoretical specification and empirical operationalization of process and context factors determining competence development and school success.
- Measures are needed that can be used to study competence development in longitudinal analyses over longer periods of time.
- Similarly, measures that are sufficiently sensitive to change are required in order to estimate the effects of interventions.
- Statistical methods need to be refined or developed for capturing change in data over time, not only at the individual level but also at the levels of schools and school systems.

- Technology-based assessment systems are needed to allow for the use of more complex and innovative test formats and, in the long run, to reduce the cost of testing.

## References:

- Baumert, J./Brunner, M./Lüdtke, O. and Trautwein, U. (2007): Was messen internationale Schulleistungsstudien? – Resultate kumulativer Wissensserwerbsprozesse. Eine Antwort auf Heiner Rindermann. *Psychologische Rundschau* 58, 118-129.
- Baumert, J./Klieme, E./Neubrand, M./Prenzel, M./Schiefele, U./Schneider, W./Stanat, P./Tillmann, K.-J. and Weiß, M. (Eds.) (2001): PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. Opladen.
- Blossfeld, H.-P./Doll, J. and Schneider, T. (2008): Bildungsprozesse im Lebenslauf – Grundzüge der zukünftigen Bildungspanelstudie für die Bundesrepublik Deutschland. *Recht der Jugend und des Bildungswesens* 3, 321-328.
- Bowers, J. and Drake, K.W. (2005): EDA for HLM: Visualization when probabilistic inference fails. *Political Analysis* 13 (4), 301-326.
- Carstensen, C.H./Prenzel, M. and Baumert, J. (2008): Trendanalysen in PISA: Wie haben sich die Kompetenzen in Deutschland zwischen PISA 2000 und PISA 2006 entwickelt? In: Prenzel, M. and Baumert, J. (Eds.): Vertiefende Analysen zu PISA 2006. *Zeitschrift für Erziehungswissenschaft: Sonderheft No. 10*.
- Ditton, H. (Ed.) (2007): Kompetenzaufbau und Laufbahnen im Schulsystem. Eine Längsschnittuntersuchung an Grundschulen. Münster.
- Döbert, H. and Sroka, W. (Eds.) (2004): Features of successful school systems. A comparison of schooling in six countries. Münster.
- Hetmeier, H. and Leidel, M. (2007): Internationale Praktiken in der Bildungsstatistik. Presentation at a workshop of the Standing Conference of the Ministers of Education and Cultural Affairs of the Länder in the Federal Republic of Germany. Berlin.
- Holtappels, H.G./Klieme, E./Rauschenbach, T./Stecher, L. (Eds.) (2008): Ganztagschule in Deutschland. Ergebnisse der Ausgangserhebung der "Studie zur Entwicklung von Ganztagschulen" (StEG). Studien zur ganztägigen Bildung, Band 1. Weinheim.
- Hosenfeld, I. (2008): Statewide monitoring and assurance of school quality: Principles of assessment and internet-based feedback of test results. In: Hartig, J./Klieme, E. and Leutner, D. (Eds.): Assessment of competencies in educational contexts – State of the art and future prospects. Göttingen.
- IQB (2008). Aufgaben und Verfahrensweise des Forschungsdatenzentrums (FDZ) retrieved on March 4, 2009 from [http://www.iqb.hu-berlin.de/arbbereiche/fdz/dateien/FDZ\\_VO.pdf](http://www.iqb.hu-berlin.de/arbbereiche/fdz/dateien/FDZ_VO.pdf). [Last visited: 02/12/2010].
- IQB (2009): FDZ. Forschungsdatenzentrum am IQB (Research Data Center), retrieved on March 4, 2009 from [http://www.iqb.hu-berlin.de/arbbereiche/fdz/dateien/FDZ\\_VO.pdf](http://www.iqb.hu-berlin.de/arbbereiche/fdz/dateien/FDZ_VO.pdf). [Last visited: 02/12/2010].
- Kanning, U.P. (2003): Diagnostik sozialer Kompetenzen. *Kompandien Psychologische Diagnostik*, Vol. 4. Göttingen.
- Klieme, E./Helmke, A./Lehmann, R.H./Nold, G./Rolff, H.-G./Schröder, K./Thomé, G. and Willenberg, H. (Eds.) (2008): Unterricht und Kompetenzerwerb in Deutsch und Englisch. Weinheim.

- Klieme, E. and Steinert, B. (2008): Schulentwicklung im Längsschnitt. Ein Forschungsprogramm und erste explorative Analysen. *Zeitschrift für Erziehungswissenschaft*, Sonderheft 10/2008, 221-238.
- Köller, O. (2008): Bildungsstandards – Verfahren und Kriterien bei der Entwicklung von Messinstrumenten. *Zeitschrift für Pädagogik* 54, 163-173.
- Konsortium Bildungsberichterstattung (2006): Zur langfristigen Sicherstellung der Datenbasis für die Bildungsberichterstattung. Deutsches Institut für Internationale Pädagogische Forschung (DIPF). Frankfurt am Main.
- Konsortium Bildungsberichterstattung (2007): Das weiterentwickelte Indikatorenkonzept der Bildungsberichterstattung. Deutsches Institut für Internationale Pädagogische Forschung (DIPF). Frankfurt am Main.
- Kunter, M. and Stanat, P. (2002): Soziale Kompetenz von Schülerinnen und Schülern: Die Rolle von Schulmerkmalen für die Vorhersage ausgewählter Aspekte. *Zeitschrift für Erziehungswissenschaft* 1, 49-71.
- OECD (2001): Knowledge and skills for life. First results from PISA 2000. Paris.
- Prenzel, M./Baumert, J./Blum, W./Lehmann, R./Leutner, D./Neubrand, M./Pekrun, R./Rost, J. and Schiefele, U. (Eds.) (2006): PISA 2003. Untersuchungen zur Kompetenzentwicklung im Verlauf eines Schuljahres. Münster.
- Seidel, T./Prenzel, M./Rimmele, R./Herweg, C./Kobarg, M./Schwindt, K. and Dalehefte, I.M. (2007): Science teaching and learning in German physics classrooms – Findings from the IPN-Video Study. In: Prenzel, M. (Ed.): Studies on the educational quality of schools. The final report on the DFG Priority Programme. Münster.
- Stanat, P. and Lüdtke, O. (2008): Multilevel issues in international large-scale assessment studies on student performance. In: van de Vijver, F.J.R./van Hemert, D.A. and Poortinga, Y.H. (Eds.): Individuals and cultures in multilevel analysis. Hillsdale.
- Stanat, P. and Lüdtke, O. (2007): Internationale Schulleistungsvergleiche. In: Trommsdorff, G. and Kornadt, H.-J. (Eds.): Enzyklopädie der Psychologie: Kulturvergleichende Psychologie, Band 2: Kulturelle Determinanten des Erlebens und Verhaltens. Göttingen.
- Wagner, G.G./Frick, J.R. and Schupp, J. (2007): The German Socio-Economic Panel Study (SOEP): Scope, evolution and enhancements. *Schmollers Jahrbuch* 127 (1), 139-170.



### 3.4 Knowing More about Vocational Training

*Steffen Hillmert*

Contact:

Steffen Hillmert  
University of Tübingen  
Department of Sociology  
Wilhelmstr. 36  
72076 Tübingen  
Germany  
e-mail: [steffen.hillmert\[at\]uni-tuebingen.de](mailto:steffen.hillmert[at]uni-tuebingen.de)



## Abstract

Modern societies depend on the successful and comprehensive provision of skills. In younger cohorts, the majority of the population has received some form of vocational training, increasing the demand for timely information about the various forms of training and their relation to the broader societal context. Over recent decades, the patterns of participation in education and training have become more complex and heterogeneous, and extend further into the life course. In considering this development, this paper discusses the extent to which existing and projected data sources are suitable for investigating scientific and policy-related questions in this field. Among these questions are: what trends in vocational training can be identified over an individual's lifespan? What are the relative chances of receiving specific types of training? Who, in particular, is likely to receive the most attractive types? Are training measures effective? When reviewing currently available data, it becomes clear that progress has been made in the past few years. It is also obvious, however, that fundamental questions cannot presently be answered on the basis of the available large-scale data on vocational education and training. Some key recommendations are presented to remedy these gaps.

Keywords: vocational training, data, research infrastructure, overview, Germany

## 1. Analytical framework and research questions

Modern societies depend on the successful and comprehensive provision of skills. In younger cohorts, the majority of the population has received some form of vocational training, increasing the demand for timely information about the various forms of training and their relation to the broader societal context. This has raised questions like: what do participation rates in training currently look like? What are the relative chances of receiving specific types of training? Which groups are likely to receive the most attractive types? Are training measures effective?

Over recent decades, patterns of participation in education and training have become more complex and heterogeneous. Young people do not necessarily receive training from just one type of institution, but combine episodes of training in various ways (Hillmert and Jacob 2003). It therefore makes sense to speak of educational and training *careers*, which extend ever further into the life course. Today, many individuals participate in formal education for more than two decades. For research, this highlights the need for studying education and training from a dynamic, lifelong perspective based on empirically observed behavior, not administrative categories. Accordingly, the demand for detailed and reliable data is high, for such information provides a solid empirical foundation for future educational

policies. While it is still important to know about aggregate numbers (such as training positions, participants, or applicants at a particular point in time), it is crucial to look at individual situations and the dynamic processes that take place within education and training. This means that it is essential to collect not only “snapshot” information about current activities in vocational training, but also information about connections with the individual’s previous life history and subsequent career steps.

In Germany, recent and ongoing longitudinal studies (e.g., SOEP,<sup>1</sup> GLHS,<sup>2</sup> BIJU,<sup>3</sup> LAU,<sup>4</sup> ULME,<sup>5</sup> BIBB<sup>6</sup> Transition Study, DJI<sup>7</sup> Transition Panel) have compiled valuable data about educational careers – sometimes including measures of performance of competencies – and have thus enhanced our knowledge about likely causal relationships. Nevertheless, evidence-based policies continuously require differentiated, reliable, and up-to-date information. It is therefore necessary to improve the quality of process-produced data in line with new demands. Importantly, however, it is necessary to keep in mind that official statistics are constructed on a specific legal basis, so any recommendations for improvement have to take into consideration the details of these regulations. This may pose problems for the collection of comprehensive data on the training situation of the whole population, as this would require still more radical innovations in the information infrastructure such as the introduction of a population register. Even when ambitions are lower, there still exists an urgent need for improving the current data situation. Even a brief review reveals that very basic information about the situation of vocational training in Germany is not available. Measures to change this would often require no more resources yet still be of great value. In any case, research will continue to operate with a variety of different data sources, which should be coordinated as much as possible.

To overcome the aforementioned problems, a simplified framework should be used, containing basic requirements for the systematic collection of information on training. The initial focus of any approach proposing to assess the performance of the vocational education and training (VET) system should be to differentiate between its institutional characteristics; that is, em-

---

1 German Socio-Economic Panel (SOEP, *Sozio-oekonomisches Panel*).

2 German Life History Study.

3 Learning Processes, Educational Careers, and Psychosocial Development in Adolescence and Young Adulthood (BIJU, *Bildungsprozesse, Kompetenzentwicklungen und Selektionsentscheidungen im Vor- und Grundschulalter*).

4 Aspects of Learning Prerequisites and Learning Development (LAU, *Aspekte der Lernausgangslage und der Lernentwicklung*).

5 Study on Achievement, Motivation, and Attitudes of Students at the Beginning of Vocational Education (ULME, *Untersuchungen der Leistungen, Motivation und Einstellung zu Beginn der beruflichen Ausbildung*).

6 Federal Institute for Vocational Education and Training (BIBB, *Bundesinstitut für Berufsbildung*).

7 German Youth Institute (DJI, *Deutsches Jugendinstitut*).

ploy an “institution-oriented differentiation” approach. Historically, vocational training statistics have been equated with information on apprenticeships, but today it seems useful to distinguish among three major areas of the VET system below the tertiary level (Baethge et al. 2003; Autorengruppe Bildungsberichterstattung 2008): the “dual system,” i.e., apprenticeships organized as a combination of firm-based training and vocational schooling; (full-time) “school-based training”; and also a large “transition system” of measures including youth training schemes and basic forms of vocational training. Additionally, there exist special cases like training in the Civil Service.

A second issue to be considered when assessing the VET system is that one can differentiate among different analytical elements. These elements, also known as central “dependent variables” for empirical analyses, represent the basic building blocks of careers in education and training. They allow inferences to be made about mobility and developments in the training system. These elements relate to several different concerns, in particular, to

- access: the transitions to specific forms of vocational training and their determinants;
- outcomes: learning, competencies, and qualifications in vocational training; and
- impact: in the sense of transitions out of vocational training and their consequences, in particular labor market consequences.

Complementing the institution-oriented differentiation approach, and just as necessary, is an approach that considers non-institutional characteristics; that is, an “individual-level differentiation” approach. This approach allows researchers to compare the relative chances of particular groups in society to achieve particular educational outcomes and impacts. For empirical analyses, this approach would focus on the set of central “independent variables.” These analytical dimensions are summarized in Table 1. While variables like age and gender are commonly used distinctions, differences in nationality and – even more important for meaningful analyses – migration background have only recently come into the focus of official statistics. Another important form of non-institutional differentiation is regional differentiation. Classifying various aggregate units may extend the database further towards a multilevel structure.

It is essential for the analysis of training careers to make use of longitudinal data. For research, *prospective* collection of information would be preferable, i.e., following the careers of individuals as they develop over time. However, such data designs have tended to be controversial with respect to data protection as they necessarily require matching information on individual cases across several waves of data collection. Hence, the second-best solution is collecting information on individual developments *retro-*

*spectively*, i.e., gathering time-referenced information about individuals' previous experiences and actions each time someone begins a (new) period of activities in VET.

This framework is rather simple and selective. When contrasted with the basic questions in the first paragraph, however, it does make clear that fundamental and highly relevant questions cannot presently be answered on the basis of the currently available large-scale data on VET. These questions have already yielded significant findings in small-scale studies and hence warrant further investigation. Among these questions are: how many people combine two or more episodes of VET? Do young people with and without a migration background have similar chances of access to training? How do people with a migration background perform within the system? Does training within the dual system and in full-time vocational schools lead to similar patterns of transitions to employment? Do training programs result in transitions to regular training and/or stable employment? What proportion of a cohort entering vocational training completes it successfully?

Table 1: Systematic information for a dynamic analysis of vocational training

<b>Institution-oriented differentiation</b>	<b>Analytical elements in a life-course perspective</b>	<b>Individual-level and other non-institutional differentiation (examples)</b>
<ul style="list-style-type: none"> <li>- Dual system</li> <li>- School-based training</li> <li>- Transition system</li> </ul>	<ul style="list-style-type: none"> <li>- Access</li> <li>- Outcome/performance</li> <li>- Impact/consequences</li> </ul>	<ul style="list-style-type: none"> <li>- Individual history of education and training</li> <li>- Age</li> <li>- Gender</li> <li>- Migration background</li> <li>- Regional differentiation</li> </ul>

## 2. Contrasting demands and available databases

This section discusses the extent to which existing and projected data sources can be used to investigate such questions. Compared to other industrialized countries, VET in Germany is formalized to a relatively high degree. Due to the differentiation and the complex institutional structure of the German education and training system, however, relevant information sources are very heterogeneous. Given this, any brief overview of relevant data and their characteristics – taking into account the basic distinctions between the three sectors of the VET system – is necessarily a selection (for a general overview

on the information infrastructure in this area, see Weishaupt and Fickermann 2000; Baethge et al. 2003, 47-54; Bellmann 2005; Brosi 2005).

### *2.1 Data on vocational training within the Dual System*

Historically, the data situation regarding apprenticeship training has been adequate. Given its dual nature, it is covered by various data sources. These include the School Statistics, which are collected by the *Länder*, and the Vocational Training Statistics of the responsible authorities, i.e., institutions such as the Chamber of Commerce, whose statistics are often referred to as *the Vocational Training Statistics*. Additionally, there are also statistics on the vocational training market that report the number of registered applicants and vacancies. These allow for calculations of supply and demand, at least in a simple form (BMBF 2007). Additional information about apprenticeships is provided by Employment Statistics and regular collections of firm-level data (IAB Establishment Panel).

The Vocational Training Act (BBiG, *Berufsbildungsgesetz*) is the central legal basis of the VET Statistics. As this is national law, it allows for standardized gathering of information. The Reform Act of 2007 has significantly improved the collection of mandatory data, in particular in terms of the shift from aggregate to individual-level reporting and the provision of more information on individual-level determinants of transition (Uhly 2006). Implementation of these changes is, however, still underway.

Despite some level of progress, there still exist considerable deficits for dynamic analyses: there is no linkage between years and no information on timing. In other words, there is information on transitions within a given year, but it represents a one-off snapshot. This means that there is insufficient information about the length of training careers, the structure of multiple training episodes, and in general on lifelong aspects. Since program termination rates are calculated on a yearly basis, it is not possible to distinguish between temporary and final dropout and hence to calculate cohort-specific rates of success within the apprenticeship system.

*This situation could easily be remedied by the introduction of a personal ID number which would allow for connecting individual-specific data entries across different years.* Currently, whilst information on previous education and training – including school-based training and qualification measures – is indeed collected (Statistisches Bundesamt 2007a), this information is restricted to type and level. Furthermore, the range of these statistics is limited. An implication of this is that matching person-specific data records across years (and institutions), which is the preferable solution for reconstructing training biographies, is not possible. As long as this cannot be done, it would be helpful to collect qualitative information on previous VET experience (most notably, occupation) and on the *timing* of previous experiences in training.

Another specific deficit with respect to currently available large-scale data on vocational education and training is the insufficient amount of information about possible migration backgrounds. This additionally required data needs to go beyond information about the individual's nationality. It would be helpful to coordinate the definitions of an extended set of collected variables with other central data sources in this area, such as the individual-level School Statistics, where information is made available about country of birth, parents' country of birth, and year of immigration, as well as language spoken at home. Including at least some proxy information on regional mobility among apprentices or previous performance in the form of grades may also be considered.

## 2.2 Data on full-time school-based training

The varying forms of full-time school-based training have become an increasingly important part of the VET system. They form a mixed category, but a major share consists of training in occupations outside the BBiG regulations. Whereas vocational schools are regulated by state law and covered by the School Statistics (Statistisches Bundesamt 2007b), data are collected by the *Länder* according to varying regulations and classifications. In many respects, the information collected and published has been limited and heterogeneous (Krüger 2005).

A fundamental systematic deficit of statistics that focus on secondary school students is that there is no information about applicants or the amount of space available in school programs. This means that – in contrast to apprenticeships – there is no information about relations of supply and demand in this sector of VET. Special problems are associated with data on school-based training in the healthcare sector. Data are incomplete (e.g., there is no reporting in the federal state of Hesse) and of varying quality. One of the reasons for this is that supervision lies with different authorities, some of them with no obligation to report.

A consequence of all this is that there is still no comprehensive account of the volume and the structure of school-based training in Germany. Given this heterogeneity and limitations, major improvements can be expected from the coordinated “core data” project on schools (KDS, *Kerndatensatz für schulstatistische Individualdaten der Länder gemäß dem Beschluss der Kultusministerkonferenz*) recently undertaken by the *Länder* – that is, as long as there is comprehensive coverage of all vocational schools in all federal states. Similar to the changes in the BBiG statistics, this undertaking would also mean a shift from aggregate reporting to the collection of individual-level information and an extension of collected variables. For this change to be of use for research on training careers, it is essential that individual records be able to be matched from one year to another. However,

implementation of this measure has caused some controversy and has so far proceeded very differently in the various *Länder*.

### 2.3 Data on the “transition system” of training measures

The term “transition system” subsumes a variety of measures within the VET system, including both youth training schemes and forms of preparatory vocational training (including attainment of general school qualifications). It is a very heterogeneous category, but given its considerable size, it makes sense to include it in any regular assessment of the VET system. So far, it is not clear how permanent and successful the measures included in the transition system are as components of individual training careers. Empirical studies have hitherto relied on one-off samples or focused on special, temporary programs (Troltsch et al. 1999; Dietrich et al. 2002). For a specific sub-group of the training population, there is analytical potential in matching different sources of process-produced data on training and employment. Notably, this undertaking is an ongoing project, entitled “Integrated Employment Biographies,” at the Institute for Employment Research (IAB, *Institut für Arbeitsmarkt- und Berufsforschung*).

Regular reporting about transition system measures is mainly based on the statistics of vocational schools and the statistics of measures financed by the Federal Employment Agency (BA, *Bundesagentur für Arbeit*). In terms of these reported figures, it is difficult to identify any overlap (Autorengruppe Bildungsberichterstattung 2008: 99). Clearly, there exists a need for better coordination between the two data sources. While there is some information available on transitions to employment after the end of particular measures (“integration into the labor market”), there is no comprehensive and permanent data source available that records intertemporal individual experiences within the transition system and links them to other forms of education and training.

The lack of transparency regarding empirical information about the transition system has further consequences. For example, as young people in these measures do not appear as applicants in other sectors, it is very difficult to assess the overall demand in the VET system. For this part of VET, it is especially important to know more about the incidence, duration, and success of training measures undertaken by individuals. To overcome the shortcoming that information across various years of data collection cannot be matched, information on previous experiences in the transition system needs to be collected when individuals enter training in either the dual system, school-based training or any another training measure. In order to relate this information to the relevant groupings of former participants in these training sectors, however, it would be most important to have infor-

mation on the *timing* of these experiences. Again, it would also be essential to collect information on the individual migration background.

#### *2.4 Additional data sources*

Apart from data that have been designed primarily to inform about the VET system, there exist a number of other representative data sources that do not have such a specific focus, but nevertheless provide relevant information. The most prominent of these examples is the German Microcensus, which is a valuable source of information on the distribution of qualifications differentiated by individual characteristics. In contrast to many other data sources, recent releases of the Microcensus contain fairly comprehensive information about both nationality and migration background. Although conceptualized as a short-term panel, in principle the Microcensus is able to map educational transitions. Notably, there are a few direct indicators in the questionnaire that could be extended. For example, as of 2005, information on additional vocational degrees is available; however, this is restricted to graduates from higher education so it does not allow for the identification of multiple training episodes in secondary vocational training. In relation to additional data sources, such as the Microcensus, refining and harmonizing the definitions of specific variables may greatly enhance their compatibility and hence their value for empirical research on training issues. In general, conceptualizing and further developing these data sources should be done in close cooperation with potential users. One good way to accomplish this would be to conduct issue-specific expert workshops.

### **3. Questions of data access and use**

Collecting adequate data is crucial, but for practical research, data availability and access are also key issues. In recent years, major progress has been made regarding the systematic and regular collection of individual-level data on the VET system. There is still a great need for well-regulated access to and systematic documentation of these data. However, up to this point it is not clear to what extent and in what form this information will be accessible to scientific researchers.

Other major advances can be expected from the recently established Research Data Center for vocational training data at the BIBB. Given that relevant data on the VET system are held by very different institutions, a common and regularly updated directory would be helpful for researchers.



Another concern regarding data access and use is the range of available databases beyond official statistics. While data access and support for established surveys like the SOEP are exemplary, many issue-specific datasets never become available or even known to potential users because the principal investigators are obliged by their funding contract to delete the data once the project is finished. In light of this, many researchers argue that it is crucial that the data collected for projects commissioned by public authorities be made available through the relevant data archive or Research Data Center.

#### **4. Summary and recommendations**

Contemporary, individualized knowledge-based societies require comprehensive, up-to-date, and dynamic data, that is, regularly collected data on individual histories of education and training rather than just aggregate snapshots. Taking this into consideration, the current data situation concerning the VET system can be summarized as follows:

Information about the dual system of apprenticeship training is fairly comprehensive, but the shift towards an individual-level statistical accounting system should be complemented by a systematic decision allowing for the study of longer sequences of training careers rather than single transitions. The simplest solution for this would be a permanent ID number.

In contrast to the data on the dual system, the data available on full-time vocational training is much more limited. This situation is unsatisfactory, not least in terms of the gender-specific participation in these institutions. It is also fairly heterogeneous due to the federalist structure of the German school system. Further efforts of coordination are necessary to build up a regular accounting system that allows dynamic analyses on an individual level. A further problem exists in that there is still no systematic reporting on transition system measures that may be related to transitions to regular forms of training.

In spite of the progress that has been made, a major deficit remains: all three main sectors of the VET system are covered by very different regulations and procedures for data collection. This means that transitions between these sectors can be analyzed only in a very selective and limited manner. Moreover, it is impossible to use process-produced data to study some very interesting research questions about the links between the VET system and other educational institutions as they are represented in patterns of educational careers. Nonetheless, it has been established in survey research that careers in education and training do transcend the borders of specific institutions and educational levels. Linking databases on different areas of education and training therefore remains a top future priority. Again, the intro-

duction of a common ID number – in combination with a consistent concept of data protection and data availability – could help greatly in overcoming this deficit. Moreover, great care should be taken to harmonize the definitions and rules of data collection in the various sectors of the education and training system.

Considering the aforementioned shortcomings of currently available large-scale data on the VET system, there are major opportunities to improve the scope and the quality of regularly collected, process-based data. Nevertheless, despite the need for a shift towards more individualized data, large-scale surveys will continue to be indispensable and therefore still require adequate funding. Both aggregate and individualized data are necessary for studies of specific issues and causal relations. For rigorous empirical analyses of the VET system, a broader range of individual characteristics, including information about parents and families, are required. This is because educational careers are embedded in social relations, and the impact of families and life situations on educational decisions is strong. Moreover, it is necessary to include all three sectors, including the transition system, into systematic data collection, for example, in the multi-cohort National Educational Panel Study (NEPS), which is due to start in 2009/2010. Given its sample size and scope, as well as the comprehensiveness of its data, the NEPS can be expected to accomplish an unprecedented level of integration among analyses of various institutional stages of educational and training careers.

The expected trend towards international comparisons in all areas of education and training places extra pressure on the need for adequate data. These data requirements range from large-scale cross-sectional reporting and longitudinal studies to evaluation studies of specific institutions or measures. VET systems have become part of such endeavors at a relatively late stage, but large-scale assessments comparable to the Programme for International Student Assessment (PISA) (VET-LSA: Baethge et al. 2006; PIAAC, Programme for the International Assessment of Adult Competencies) have been announced. Coordination among various programs assessing both general education as well as vocational and academic training may provide new challenges. Cross-national research has shown that it is often more salient to compare whole systems of education and training than specific elements defined on the basis of nominal institutional classifications. To allow such a system evaluation, it is again important to link data from different parts of the education and training system and to harmonize the definitions used when collecting these data – on both the national and the international level.

## References:

- Autorengruppe Bildungsberichterstattung (2008): Bildung in Deutschland 2008. Ein indikatorengestützter Bericht mit einer Analyse zu Übergängen im Anschluss an den Sekundarbereich I. Bielefeld.
- Baethge, M./Buss, K.-P. and Lanfer, C. (2003): Konzeptionelle Grundlagen für einen Nationalen Bildungsbericht: Berufliche Bildung und Weiterbildung / Lebenslanges Lernen. Bonn/Berlin.
- Baethge, M./Buss, K.-P. and Lanfer, C. (Eds.) (2005): Expertisen zu den konzeptionellen Grundlagen für einen Nationalen Bildungsbericht. Bonn/Berlin.
- Baethge, M./Achtenhagen, F./Arends, L./Babic, E./Baethge-Kinsky, V. and Weber, S. (2006): Berufsbildungs-PISA: Machbarkeitsstudie. Stuttgart.
- Bellmann, L. (2005): Der Stand der Aus- und Weiterbildungsstatistik in Deutschland. In: Baethge, M./Buss, K.-P. and Lanfer, C. (Eds.): Expertisen zu den konzeptionellen Grundlagen für einen Nationalen Bildungsbericht. Bonn/Berlin.
- Brosi, W. (2005): Anmerkungen zur Verfügbarkeit von Statistiken im Bereich der beruflichen Bildung als Basis für eine umfassende Bildungsberichterstattung. In: Baethge, M./Buss, K.-P. and Lanfer, C. (Eds.): Expertisen zu den konzeptionellen Grundlagen für einen Nationalen Bildungsbericht. Bonn/Berlin.
- Bundesministerium für Bildung und Forschung (BMBF) (2007): Berufsbildungsbericht 2007. Berlin.
- Dietrich, H./Behle, H./Böhm, R./Eigenhüller, L. and Rothe, Th. (2002): Design der IAB-Begleitforschung zum Jugendsofortprogramm im Kontext des IAB-Forschungsschwerpunkts School-to-Work-Transition. Nuremberg.
- Hillmert, S. and Jacob, M. (2003): Bildungsprozesse zwischen Diskontinuität und Karriere: das Phänomen der Mehrfachausbildungen. Zeitschrift für Soziologie 32 (4), 325-345.
- Krüger, H.(2005): Zur Datenlage vollzeitschulischer Berufsausbildung. In: Baethge, M./Buss, K.-P. and Lanfer, C. (Eds.): Expertisen zu den konzeptionellen Grundlagen für einen Nationalen Bildungsbericht. Bonn/Berlin.
- Statistisches Bundesamt (2007a): Berufsbildungsstatistik – Begriffe und Erläuterungen. Wiesbaden.
- Statistisches Bundesamt (2007b): Fachserie 11, Reihe 2: Bildung und Kultur – Berufliche Schulen. Wiesbaden.
- Troltsch, K./László, A./Bardeleben, R. von and Ulrich, J.G. (1999): Jugendliche ohne Berufsausbildung. Eine BiBB/EMNID-Untersuchung. Bonn.
- Uhly, A. (2006): Weitreichende Verbesserungen der Berufsbildungsstatistik ab April 2007. In: Krekel, E.M./Uhly, A. and Ulrich, J.G. (Eds.): Forschung im Spannungsfeld konkurrierender Interessen. Die Ausbildungsstatistik und ihr Beitrag für Praxis, Politik und Wissenschaft. Bonn.
- Weishaupt, H. and Fickermann, D. (2000): Informationelle Infrastruktur im Bereich Bildung und Kultur, Expertise für die Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik. Erfurt.



## 3.5 Higher Education

*Andrä Wolter*

Contact:

Andrä Wolter  
Dresden University of Technology  
Faculty of Education  
01062 Dresden  
Germany  
e-mail: [andrae.wolter\[at\]tu-dresden.de](mailto:andrae.wolter[at]tu-dresden.de)

## **Abstract**

During the last five years, higher education research in Germany seems to have taken a significant upturn. This is partly a side effect of the obvious boom overall in empirical research on education, and partly of the reform movement that has affected the German higher education system since the mid-1990s. The demand for data in the field of higher education will increase considerably in the future. The available data infrastructure for higher education research in Germany consists of two central and complementary sources: official statistics on higher education on the one hand and survey-based research on the other.

In principle, there are no serious obstacles to accessing the available data stock relevant to higher education. However, access to some of the most important surveys would be improved through the establishment of a Research Data Center at the Higher Education Information System (HIS, *Hochschul-Informationen-System*). Furthermore, there are some significant shortcomings in the present data provision. New topics that place new demands on data provision must be integrated into official statistics and survey-based research alike (e.g., migration status, competencies, lifelong learning, quality of studies, institutional effects, international mobility, programs to promote younger scholars etc.). In particular there is a lack of panel designs. The recently established National Education Panel Study (NEPS) will eliminate some, but not all, of these deficiencies.

## **1. The development of higher education research in Germany: Old and new research questions**

In Germany, as in other countries, higher education is faced with increasing pressure to justify its existence in terms of demand, outcomes, effectiveness, study success, and other issues that might be grouped collectively under the concept of accountability. Researchers and particularly politicians – at institutional, state, and national levels – are interested in information and data concerning the results and performance of higher education institutions, citing funding bottlenecks and growing competition for students, scholars, academic reputation, and resources. With the growing social and economic centrality of higher education, the academic and political interest in data on the development and functioning of its institutions is rapidly increasing.

Over the last five decades a wide range of academic and political issues in higher education have been the focus of empirical and non-empirical higher education research in Germany. Although higher education had already been the subject of various academic undertakings since the late 1950s, research on higher education in general remained on the periphery of social and educational research for a long period of time. As a result of the massive quantitative expansion of higher education over the past few

decades, colleges and universities have become some of the most important institutions – sometimes regarded as the most important institution of professional education in modern, knowledge-based societies. In view of the overall trend toward upgrading the qualification structures of employment, higher education can be expected to become even more relevant as an institutional center of qualification. As a result, the academic and political interest in higher education can be expected also to increase considerably.

During the last five years, higher education research in Germany seems to have taken a significant upturn. This is partly a side effect of an obvious boom overall in empirical research on education, and partly a result of the reform movement that has affected the higher education system in Germany since the mid-1990s. As with empirical educational research in general, research on higher education seems to have profited from the new paradigm of evidence-based educational policy. One element of this is the establishment of continuous monitoring systems at different levels (international, national, state), including higher education, such as the German National Report on Education (*Nationaler Bildungsbericht*) (Avenarius et al. 2006; Klieme et al. 2008), which rapidly gave reinforcement to the need for an elaborated data infrastructure. The fact that the higher education system in Germany, as in other countries, is currently the subject of a lively public reform debate and faces several reform challenges (Wolter 2004; 2007a), has also stimulated the increasing interest in higher education research.

Higher education research covers a broad range of research questions and topics and it is difficult to reduce the diverse research activities that have been pursued in recent years to a selection of a few major themes. However, it may be possible to distinguish the following four main fields of research (Teichler 2002):<sup>1</sup>

- *Quantitative-structural changes in higher education.* This area of inquiry considers the development of the social demand for higher education, the consequences of the wholesale “massification” of the system, the particular institutions, the institutional structures of higher education and its changes (e.g., through diversification, profiling, or vertical and horizontal differentiation), the provision of studies, the interdependencies between the expansion, types of differentiation, etc.;
- *Transitions and processes of studies.* A significant amount of research has focused on topics dealing with the first transition point – the status passage between school and higher education, including issues such as access and admission, social inequality and opportunities to study, the social and economic conditions of studying, processes and success of

---

1 Because research is the second pillar of higher education, another important field is represented by research on research or on science, the subject of another article in this volume by Stefan Hornbostel.



studying (including the drop-out phenomenon) and their determinants, teaching and learning, and student mobility. Also included under this heading are issues arising from a second transition point – from higher education to the labor market and employment, including early vocational careers, the match between higher education and employment, and similar questions;

- *Post-graduate training and academic staffing.* This field encompasses topics such as the different stages and paths to a professorship, the effectiveness and quality of doctoral programs, the main activities and time budget of the academic staff, faculty development, employment conditions, and the career perspectives, especially of young scholars, etc.;
- *Organization, management, and governance of higher education.* This field includes, at the systemic or institutional level, topics such as the external relationships between state and university, the internal organization of institutions, issues of efficiency, funding, professional institutional management, new concepts and procedures of steering and allocation, evaluation and quality assurance, etc.

The kinds of data required in these studies vary with the research questions and fields. Whereas the first three areas require primarily data at the national level, most research on steering and governance topics depends on the availability of data primarily at an institutional level. The following discussion focuses on available data stocks and new data requirements primarily at the national, rather than at the institutional level. For this reason, the new wave of ranking procedures with their enormous demand for differentiated data on the performances of individual institutions of higher education (Bayer 2004; Statistisches Bundesamt 2007) is not the subject of this report.

Sometimes it is difficult to differentiate strictly between old and new research questions. Old questions often remain relevant over time or become significant again in changing contexts. For example, the effect of the continuous expansion of participation in higher education on the changing relationships between higher qualifications, profession, and employment had already become a hotly debated issue in the early 1970s. But these issues have become relevant again with the new wave of expansion that has occurred since the 1990s. They will retain their relevance with the recent political consensus that higher education in Germany should aspire to a participation rate of 40 percent of the corresponding age group and a graduate rate of 35 percent (Wissenschaftsrat 2006) in order to keep pace with other highly developed countries. Therefore, issues such as the unemployment of graduates or whether there is an adequacy link between qualifications and employment have always been acute questions, even if now they are being analyzed in a methodologically more differentiated way and in the changed context of the rising knowledge-based society.

On the other hand, there are new questions and topics that have arisen in the context of these enormous changes to German higher education – which might even be called a fundamental transformation – during the last decade:

- *The impact of the Bologna Process.* During the last few years interest has grown considerably in the impact of the Bologna Process on studies and studying (Teichler 2008; ZSE 2008). At a structural level, changes in the relationships between different types of institutions (e.g., trends of convergence) and in the provision and organization of courses have been the focus of this interest. But the Bologna accords include more than new degrees and a consecutive structuring of studies. This notion indicates a radical change in educational culture that includes changes in teaching and learning styles. So, at an individual level, the motivation to study and the learning behavior of the students and their adaptation to the new configuration of studies is of great interest. This will result in an expanding demand for related data particularly from research on students.

Since the Bologna Process has determined that employability – whatever this concept may mean (Schaeper and Wolter 2008) – is a primary objective of study, further research will also focus on transitions between higher education and work. Thus, there is (and will be) a demonstrable and increasing interest in graduate studies, manifest in cross-sectional or, more productive yet, as panel studies (HRK 2007). Furthermore, research will be concerned with the outcomes of studying, especially in terms of disciplinary competencies and broader, transdisciplinary “key competencies” (Schaeper 2005; Schaeper and Spangenberg 2007). Last but not least, the line differentiating initial and continuing studies has become blurred under the Bologna Process, which has strengthened the focus on lifelong learning as the primary mission of higher education. But the situation of continuing higher education is particularly problematic, with completely disparate and inadequate data provision (Wolter 2007b).

- *Internationalization of higher education.* The internationalization of higher education – due to globalization, Europeanization, or other processes (Teichler 2004) – has led to a growing need for internationally comparative data. Institutions of higher education today are often considered to be organizations acting on a global market with global competition. Research on higher education must take this international or even global character into account. During the last decade, this internationally comparative perspective has been reinforced for two major reasons. First, there is a political demand for comparative analyses that reveal the strengths and weaknesses of higher education systems and institutions, for information needed for the identification of concepts or models of institutional reform, and for establishing procedures of quality control. Second, the establishment of a “European Higher Education Area” in the

course of the Bologna Process has reinforced the earlier interest in student and staff mobility. As a result, there is a growing need for valid international data on higher education, in particular on student mobility (Kelo et al. 2006).

- *Differentiation and governance of higher education.* The growing stratification of the German higher education system in the course of the Excellence Program and other mechanisms of differentiation is another area of growing interest. Although these institutions of “excellence” (in the three areas of graduate schools, clusters of excellence, and future development concepts) have been selected on the basis of a state-regulated nomination procedure rather than by a market-shaped process of competition (and thus is based on reputation and performance), in the long run the claim of excellence will require academic justification by measurable criteria. This will result in an increasing need for data pertaining to the achievements of higher education institutions, primarily in the areas of research, but probably also in teaching (Hornbostel 2008a; b). Research on the results and changes, generated by the progressive implementation of new governance and steering structures, will be another important future research area (Wolter 2007a). But both of these questions are still in their infancy, since the dynamic of these changes is at an early stage. Both require, as mentioned above, special data particularly at institutional rather than the national level.

## **2. The current state of data infrastructure and the challenges in higher education research**

The available data infrastructure for higher education research in Germany consists of two main sources:

- official higher education statistics, which include information about students, personnel, and finance statistics; and
- data and results from survey-based research, in particular in the field of student and graduate research, which is conducted by research centers, such as the Higher Education System (HIS, *Hochschul-Informationssystem*) GmbH Hannover, the International Center for Higher Education Research (INCHER) at the University of Kassel, the Research Group on Higher Education (*Arbeitsgruppe Hochschulforschung*) at the University of Konstanz, and other centers.

In principle, the availability of and access to public data and survey data is ensured. Yet, with the exception of official statistics, research institutes or centers are often faced with obstacles due to lack of personnel or technical capacity. From an international perspective, higher education research may be less developed in Germany than elsewhere relative to the size of its national higher education system. However, the main problem in German higher education research does not primarily consist of a lack of data but rather a lack of an extensive and methodologically sophisticated utilization of existing data stocks. Thus, access could be improved in a practical sense, rather than through legal changes.

As a result of the higher education statistics laws, official student statistics have been presented annually by statistical offices at the state or federal level since the early 1970s. These statistics provide a significant amount of data on the number of students (new entrants or all students), their distribution over institutions and subjects, some information on their composition (gender, nationality), their regional origin, types of study entitlement, and other variables. In contrast to official school statistics, the statistics on students in higher education have consisted of individual datasets since the 1970s, so it is not necessary to establish individual statistics in this field. The official student statistics allow for many very differentiated analyses, such as on the development of (realized) student demand, on regional student mobility, duration of studies, fluctuation between subjects, and other aspects. However, there are some important limits and deficits.

- First, it has not been possible to link organized school and higher education statistics individually. So, even if it were possible to calculate general transition quotas, the transitions from grammar school to university cannot be reconstructed as individual processes. The introduction of an overarching identity number in educational statistics would therefore be an important measure that would enable the analysis of processes and transitions. This step would nonetheless confront some serious problems with respect to data protection and public acceptance.
- Second, there are many important variables that are not part of the official statistics (e.g., social origin, migrant status of students apart from formal nationality, any subjective variables). The provision of these other types of data depends completely on survey research on students.
- Third, student statistics end with exmatriculation, so of course the further life course of graduates, in particular their professional or academic careers, is not part of the student (or other) statistics. Because of this, graduate survey studies are of great importance. Some data about the employment of graduates can be gained from the German Socio-Economic Panel (SOEP, *Sozio-oekonomisches Panel*) and the Micro-

census, but these do not have the necessary depth of focus that could be obtained in graduate studies.

- Fourth, official statistics include only a very few variables (such as duration of studies) that can be used as indicators to assess the quality of studies. One of the most important gaps concerns the provision of valid and reliable data on students who drop out of programs. This is partly due to legal objections, and partly due to difficulties in the precise definition and measurement of the drop-out rate. Empirical information has primarily been generated so far by estimate models developed by HIS (Heublein et al. 2008). Currently, a joint project between the Federal Statistical Office and HIS is being carried out to deliver valid data on student drop-out.
- Fifth, official statistics do not include the universities of cooperative education (*Berufsakademien*) as a hybrid type of institution between tertiary and post-secondary education. *Berufsakademien* are not established in all German states, but where they do exist, they are often considered as the third pillar of the German higher education system, showing a high degree of curricular overlap with the universities of applied sciences, or *Fachhochschulen*.

Because of these deficits in official statistics, student survey research holds a key position in the data infrastructure of higher education research. In certain respects, official statistics on higher education and the survey-based research and data production can be seen as complementary parts of one system that shapes the research data infrastructure in the field of higher education. Survey projects can be conducted as single projects or as follow-up projects in order to build up time series. In Germany, several such follow-up studies with different target groups have been carried out since the late 1970s. They include school graduates with study entitlement, new entrants in higher education, students, and graduates.

HIS regularly undertakes various cross-sectional surveys among recent school graduates who have earned a particular study entitlement (who have an upper secondary school diploma, either the *Abitur* or *Fachhochschulreife*) and among new entrants in higher education. The focus of these studies is on the decision-making process surrounding whether to study, the choice of institution, subject of study, and the personal and social factors that determine these decisions (Heine et al. 2008a; b). The HIS survey of graduates has been partially continued as a panel during the first study sequence. On the basis of these surveys, it is possible to reconstruct the status passages between school and university as a time series for almost three decades. However, there are only a few longitudinal studies examining the complete transition process from school into higher education, starting from the upper stage of grammar school and ending at a later point in time during

university study. Some of these panel studies show a very sophisticated methodological design, but are limited to a particular *Land*, such as the TOSCA study (e.g., Köller et al. 2004).

There are two larger projects worth mentioning which concern students (across all sequences of studying) that have generally been updated every three years. The social and economic situation of students and the conditions of studying have been examined in the Social Survey (*Sozialerhebung*), also carried out by HIS since 1982 (Isserstedt et al. 2007). As a part of what is called the social dimension of the Bologna Process, a European-wide study on the social and economic conditions of studying (called Eurostudent) has been established that is also coordinated by HIS (Eurostudent 2008). The study situation, study problems, and the individual orientations of students have also been investigated since the early 1980s by the survey of students conducted at the University of Konstanz (*Konstanzer Studierendensurvey*) (Multrus et al. 2008). Some additional differentiated special analyses on certain subjects (e.g., humanities or engineering) or certain groups of students (e.g., female) have been based on this survey. Most of this type of research on students has been conducted in the framework of cross-sectional surveys.

Study on graduate education is an exploding field of research. Most of these studies focus on a retrospective assessment of studies and their outcomes – the transition from university to employment, the occupational or academic career after the first degree, and other aspects of the further life course (e.g., mobility or participation in continuing education). Since the late 1980s, there have been two research contexts on a national and international level that have provided representative data for Germany. Beginning in 1989, HIS established a study of graduates not only as a longitudinal study but also as a time series. HIS surveys a large graduate sample representative for Germany every four years with up to three panel stages: during the first, fifth, and tenth year after graduation (Briedis 2007a; b; Kerst and Schramm 2008). Coordinated by the International Center for Higher Education Research at the University of Kassel (INCHER), two internationally comparative graduate surveys have also been undertaken – CHEERS<sup>2</sup> (Schomburg and Teichler 2006; Teichler 2007a) and REFLEX<sup>3</sup> – that also embrace a larger German graduate sample and place it in a European context for comparison.

Additionally, numerous studies have been carried out at local universities or faculty levels in recent years. Many institutions are interested in the success and careers of their graduates as an indicator of academic performance or quality of studies. In the meantime, studies of graduates have also been established in three German states (Bavaria, Rhineland-Palatinate, Saxony). Graduate studies are thus one of the main areas of growing

---

2 Careers after Higher Education – A European Research Study.

3 The Flexible Professional in the Knowledge Society.

research. Nevertheless, despite or even because of the proliferation of local and regional studies, graduate surveys at the national level will retain their relevance as a benchmark for these narrower studies.

All in all, the complementary relationship between official statistics on higher education and the diverse number of surveys undertaken in this field is not a bad basis for the data infrastructure of higher education research. Nevertheless, there are some essential deficits in research and data provision.

- First, there is a lack of longitudinal studies that follow a cohort of students from upper-level grammar school (or at least from their entry into higher education) through their studies until their transition to employment and the first phase of vocational activity (with the exception of the longitudinal study of Meulemann et al. (2001) based on a cohort of grammar school graduates). Most of the existing panel studies concentrate on only one transition point – either access to higher education or to employment. This deficit is another reason why the drop-out phenomenon has not been explained sufficiently.
- Second, despite the fact there are some student surveys exploring the situation and the difficulties of students during their studies, there is a lack of data concerning the interrelationship between institutional context and the processes and outcomes of learning in higher education institutions. The relationship between internal contextual and institutional conditions at different levels (classroom, program, faculty, the institution as a whole), personal attitudes and behaviors, and learning outcomes – as well as the influence of outside learning environments – is obviously an important research desiderata. Without data on these aspects, the actual impact of institutions on learning and its outcomes is not really that clear (Pascarella and Terenzini 2005). Of course, this is an area that is both theoretically and methodologically very ambitious in light of the multiple dependent, independent, and intervening causal variables. However, it is also a venture of central importance – not only academically but also politically. The manifest trend toward more differentiation in German higher education through profiling, ranking, and excellence inevitably provokes the question of the particular influence exerted by institutions and study programs on the learning outcomes and on the later employment situations and career courses of graduates from these institutions (Teichler 2007b).
- Third, there is a considerable deficit in research and knowledge about competence development in higher education. The subject of competencies in higher education research is still a relatively new field that has become increasingly important with the Bologna Process. First of all, it is necessary to distinguish between at least three different types of competencies: (1) subject- or discipline-specific competencies; (2) cross-curricular competencies (also called “key” or “generic” competencies that

include social skills, and can normally not be acquired through learning processes independent of discipline-related learning, but which nonetheless need to be measured in their own way); and (3) competencies to act professionally in vocational demand situations. During the last ten years, there have been several attempts to measure student or graduate competencies in Germany, but these have mostly been only cross-sectional, primarily measuring for cross-curricular competencies and based largely on self-assessment or self-reporting measures (Schaeper 2005; Schaeper and Spangenberg 2007). Valid measurement procedures for competence development, based on competence tests, are very rare. In the field of discipline-related competencies, they do not even exist (apart from a very small number of pilot studies for selected subjects).

Attempts to develop and practice competence measurement procedures primarily have two main problems to confront: (1) the great diversity of discipline contexts in higher education, which is different than the school system with its core curriculum and a limited number of subjects; and (2) a completely different target group and institutional context, both of which make it more difficult to implement test-based procedures of competence measurement for students – and even more so for graduates – than is the case for pupils and classroom situations, including the acceptance of such procedures to graduates. However, there is no doubt about the relevance and necessity of the development and implementation of more elaborate procedures of competence measurement in higher education, in particular to address the question whether (or to what extent) institutions and programs actually impart the competencies they should and to what extent other formal or informal learning settings intervene in this process.

Some of these questions and issues are the subject of the recently established National Educational Panel Study (NEPS) (Blossfeld 2008; Blossfeld in this publication), which also includes a student cohort. The NEPS focuses on a number of aspects: first, on the development of competencies through higher education, mainly cross-curricular competencies; second, on the influence of institutional settings and contexts; and third, on the educational decisions and courses including the extent and conditions of success and drop-out. In the long term, the NEPS will provide empirical information and knowledge precisely in some of the deficit areas which have been specified before. However, the NEPS limits the measurement of subject-related competencies to two selected disciplines and concentrates primarily on cross-curricular competencies.

Gender problems and issues have been dealt with in many different forms in higher education research. Whereas the participation of women in higher education has continuously increased over the last decades and, in the meantime, has become higher than that of men at least at the university level,



there are still large disparities in gender participation within specific subjects. Particularly the low degree of female participation in engineering and some of the sciences has caused concern and attracted special attention. The success rate of women is higher than that of men or, the other way round, the drop-out rate is lower. Since 2000, more females than males have graduated from universities every year, and since 2003, they have graduated in greater numbers from the higher education system in general (Klieme et al. 2008, 133, 302). It seems that the future of human capital, particularly of the highly qualified workforce, depends more and more on the supply of qualified women in the labor market. Based on these trends, some new questions arise with regard to the response of the employment system. For example, transitions of female graduates, their particular employment chances and conditions, career perspectives, and the compatibility of work and family will become or remain very important issues.

Migration has been a relatively marginal issue in German higher education research so far. First of all, it is necessary to distinguish between students with migration status and internationally mobile students who do not have residential status in Germany but stay here for the purpose of their studies. The official higher education statistics register migration only in a very narrow interpretation, based on nationality. According to this definition, approximately 3 percent of all students are migrants whereas the proportion of migrants in the population is about 9 percent (Avenarius et al. 2006, 140, 273). Based on a wider (but not exhaustive) definition of migration background, including educational residents (*Bildungsinländer*), students with double nationality, and naturalized students, the proportion of migrants in the student body comes to about 8 percent compared with a proportion of this group of about 19 percent in the population (Isserstedt et al. 2007, 435). Obviously, migration has been up to now only a peripheral topic in higher education research, resulting in a lack of data and empirical knowledge despite the fact that a higher rate of participation of migrants would be a new source of social demand.

Whereas research and data provision on students and studying is relatively well-established, the state of research and data in the field of academic recruitment and academic staff is not satisfactory in the same way. Even if the official personnel statistics can deliver a lot of quantitative and structural information, there have not been any regular parallel surveys up to now – either for young scholars or for the complete academic staff. Furthermore, the official statistics cannot deliver any reliable information about the volume, the paths, and the overall situation of the new academic generation. In this area a lack of quantitative information and some important research desiderata dominate (BMBF 2008; Burkhardt et al. 2008).

Because of the current generation, changes in academic staffing, and the high demand for scholars – not only from higher education institutions but

also from the non-university research sector – the state of data provision in this area is absolutely unsatisfactory. However, some empirical, partly comparative studies have been carried out during the last years concerning the situation of young scholars and the paths of qualification and employment on the way to a professorship (Enders and Bornmann 2001; Enders and Mugabushaka 2005; Burkhart 2008; Kreckel 2008). These studies highlight the urgency of the problem. But neither the number of young scholars, currently employed at German universities as the coming generation of professors (*wissenschaftlicher Nachwuchs*), nor the number of PhD students, is precisely known. The same is true for the expanding group of postdocs.

Most quantitative information in this field is based on estimates, case studies or other limited projects. The success rate of PhD candidates is unknown as well. At best, the number, situation, and success of PhD candidates in graduate schools or with other institutions can be or has been examined, but this group represents only a small proportion. One reason for this insufficiency in data provision is the individual diversity and heterogeneity in the qualification routes, in particular to acquire a PhD degree, and in the employment conditions within and outside universities. Presently, some panel projects are being established or planned to collect more and better data on the number, routes, situation, problems, and success of this group, such as the PhD panel “ProFile” and the online panel “WinBus.” Graduate panel studies with a sufficient sample volume could be another way to improve the state of information and knowledge in this area.

### 3. Conclusions and recommendations

The demand for data in the field of higher education will increase considerably in future. This growing need is due to the rising social and political importance of higher education in postmodern societies as well as to the implementation and extension of monitoring systems including higher education. In Germany, a complementary infrastructure of research data has been being developed since the 1970s, consisting of official higher education statistics and some survey-based regular data and information sources. In the future, the National Educational Panel Study (NEPS) will significantly extend the existing system of data provision. However, there are some obvious deficiencies in the present data infrastructure, and because of this the following measures should be taken.

- *Access to data stocks:* All in all, there are in principle no serious obstacles to access to the available data stock. This is true not only for official statistics but also for the survey-based data. But access could be improved

from a practical point of view. Because HIS is the institution outside official statistics that provides the largest data stocks relevant to higher education research, a Research Data Center should be established at HIS.

- *Diversification of higher education:* In the area of official higher education, statistics on cooperative learning institutions (*Berufsakademien*) should be included in the student or personnel statistics (and also in surveys). This would take into account that fact that the structures of higher education in many countries have become blurred due to the hybrid status of some institutions – straddling post-secondary and tertiary education – and the increasing permeability between these institutions. Another important point concerns the revision of the list of disciplines in the higher education statistics, because the number and the degree of specialization within subjects have seen significant growth and provoked several serious problems of allocation.
- *Personal identity number:* The introduction of an identity number for all participants in educational programs would allow us not only to link future individual school statistics with the already given individual student statistics on pursuing transitions and additional routes of training and education, but also to improve processing data, particularly with respect to students who drop out and issues of national and international mobility. This will certainly be a delicate issue, but nevertheless is an important academic demand for the official statistics.
- *Continuation of survey-based time series:* A great deal of data and information in higher education depends on regular survey research. But this kind of survey research is based on applying for every individual project. The future availability of this data provision as the second pillar of the research data infrastructure depends completely on the continuation of these surveys. Therefore, the certainty of long-term planning is of almost constitutive importance for the data infrastructure in higher education research.
- *Indicators for quality of studies and studying:* Official statistics and survey-based research should collaborate to develop and implement a joint set of quality indicators to exploit the available data stock in a comprehensive way with regard to the increasing demand for quality assessment in higher education. Quality of studies or institutions will become one of the central issues in future higher education policy and research.
- *Longitudinal design and process data:* In student research, the most serious deficits are the lack of longer panel designs, of competence measurement, and of studies that can explore the interrelations between contextual and institutional features, personal characteristics, the pro-

cesses of studying and learning, and the learning outcomes. The NEPS will probably improve the state of knowledge in this field considerably. But longitudinal research should be intensified in general, not only in the context of NEPS. Furthermore, there should be additional pilot projects to initiate and promote the development and testing of procedures for the discipline-related measurement of competencies. Initially these should focus only on a few selected subjects, primarily those not included in the NEPS.

- *Graduate and competence studies:* Graduate studies, especially panel studies, will become even more important at all levels – at the local, state, national, and international levels. At the national level, graduate studies are indispensable as a comparative point of reference. Particular attention should be drawn to the role of institutions and programs for the allocation of position and status in the employment system. Neither the match between qualifications and employment nor, in particular, the role of competencies acquired during studies in coping with later occupational requirements are very clear. It can be expected that the trend towards more horizontal and vertical differentiation between universities will also affect the importance of institutions for employment and the future career perspectives.
- *Academic careers and young scholars:* As a part of graduate studies research, attention to the situation and further development of PhD candidates should be intensified. The lack of reliable information on the employment conditions and career paths of the younger generation of scholars, even with regard to the number of young academics or PhD candidates or to their success rate, indicates one of the most alarming deficits in the data infrastructure of higher education research. Therefore, the improvement of data provision concerning the qualification routes to an academic career remains a matter of high priority.

To sum up, it seems possible to conclude that the current state of data provision in higher education research reflects (not completely, but in many aspects) the questions, issues, definitions, and methods that emerged during the 1970s and 1980s, to which the data infrastructure has only partially adapted up to now. Many new academic or political topics and demands on data provision have arisen since this time, including issues such as migration status, competencies, lifelong learning, quality of studies, differentiation, programs to promote younger scholars, international mobility, outcomes, employability etc. These have had to be integrated into existing data programs, a venture that obviously is still ongoing. In this respect, surveys have proven to be more flexible in many respects than the often quite inflexible procedures in official statistics.

## References:

- Avenarius, H./Baethge, M./Döbert, H./Hetmeier, H.-W./Klieme, E./Meister-Scheufelen, G./Rauschenbach, T. and Wolter, A. (2006): Bildung in Deutschland. Bielefeld.
- Bayer, C.R. (2004): Hochschul-Ranking. Vorschlag eines ganzheitlichen Ranking-Verfahrens. Berlin.
- Blossfeld, H.-P. (2008): Education as a Lifelong Process. A Proposal for a National Educational Panel Study (NEPS) in Germany. 3 parts. Bamberg (unpublished).
- Bundesministerium für Bildung und Forschung (BMBF) (2008): Bundesbericht zur Förderung des wissenschaftlichen Nachwuchses. Bonn/Berlin.
- Briedis, K. (2007a): Die HIS-Absolventenstudie. In: Hochschulrektorenkonferenz (Ed.): Potentiale von Absolventenstudien für die Hochschulentwicklung. Bonn. Beiträge zur Hochschulpolitik 4.
- Briedis, K. (2007b): Übergänge und Erfahrungen nach dem Hochschulabschluss. Ergebnisse der HIS-Absolventenbefragung des Jahrgangs 2005. Hannover.
- Burkhardt, A./König, K. and Krempkow, R. (2008): Dr. Unsichtbar im Visier. Erwartungen an die Forschung zum wissenschaftlichen Nachwuchs. Die Hochschule 17 (1), 74-90.
- Burkhardt, A. (Ed.) (2008): Wagnis Wissenschaft. Akademische Karrierewege und das Fördersystem in Deutschland. Leipzig.
- Enders, J. and Bornmann, L. (2001): Karriere mit Dokortitel? Ausbildung, Berufsverlauf und Berufserfolg von Promovierten. Frankfurt/M./New York.
- Enders, J. and Mugabushaka, A.-M. (2005): Wissenschaft und Karriere. Erfahrungen und Werdegang ehemaliger Stipendiaten der DFG. Kassel.
- Eurostudent, co-ordinated by the Higher Education Information System (HIS) (2008): Social and Economic Conditions of Student Life in Europe. Bielefeld.
- Heine, C./Spangenberg, H. and Willich, J. (2008a): Studienberechtigte 2006 ein halbes Jahr nach Schulabschluss. Hannover.
- Heine, C./Krawietz, M. and Sommer, D. (2008b): Studienanfänger im Wintersemester 2006/7. HIS-Projektbericht.
- Heublein, U./Schmelzer, R. and Sommer, D. (2008): Die Entwicklung der Studienabbruchquote an den deutschen Hochschulen. HIS-Projektbericht.
- Hornbostel, S. (2008a). Exzellenz und Differenzierung. In: Kehm, B. (Ed.): Hochschule im Wandel.
- Hornbostel, S. (2008b): Evaluation der Exzellenzinitiative: Gibt es objektive Kriterien für Exzellenz? In: Bloch, R./Keller, A./Lottmann, A. and Würmann, C. (Eds.): Making Excellence. Bielefeld.
- Hochschulrektorenkonferenz (HRK) (Ed.) (2007): Potentiale von Absolventenstudien für die Hochschulentwicklung. Bonn.
- Isserstedt, W./Middendorff, E./Fabian, G. and Wolter, A. (2007): Die wirtschaftliche und soziale Lage der Studierenden in der Bundesrepublik Deutschland 2006. Bonn/Berlin.
- Kelo, M./Teichler, U. and Wächter, B. (Eds.) (2006): Eurodata. Student Mobility in European Higher Education. Bonn.
- Kerst, C. and Schramm, M. (2008): Der Absolventenjahrgang 2000/2001 fünf Jahre nach dem Hochschulabschluss. Hannover.

- Klieme, E./Baethge, M./Döbert, H./Füssel, H.-P./Hetmeier, H.-W./Rauschenbach, T./Rockmann, U. and Wolter, A. (2008): *Bildung in Deutschland*. Bielefeld.
- Köller, O./Watermann, R./Trautwein, U. and Lüdtke, O. (Eds.) (2004): *Wege zur Hochschulreife in Baden-Württemberg*. Opladen.
- Kreckel, R. (Ed.) (2008): *Zwischen Promotion und Professur*. Leipzig.
- Meulemann, H./Birkelbach, K. and Hellwig, O. (2001): *Ankunft im Erwachsenenleben. Lebenserfolg und Erfolgsdeutung in einer Kohorte ehemaliger Gymnasiasten zwischen 16 und 43*. Opladen.
- Multrus, F./Bargel, T. and Ramm, M. (2008): *Studiensituation und studentische Orientierungen. 10. Studierendensurvey an Universitäten und Fachhochschulen*. Bonn/Berlin.
- Pascarella, E.T. and Terenzini, P.T. (2005): *How College Affects Students*. San Francisco.
- Schaeper, H. (2005): *Hochschulbildung und Schlüsselkompetenzen – Der Beitrag der Hochschulforschung zur Evaluation der Qualifizierungsfunktionen und –leistungen von Hochschulen*. In: Teichler, U. and Tippelt, R. (Eds.): *Hochschullandschaft im Wandel. Zeitschrift für Pädagogik*. 50. Beiheft. Weinheim/Basel.
- Schaeper, H. and Spangenberg, H. (2007): *Die Erfassung von Schlüsselkompetenzen in den HIS-Studienberechtigten- und Absolventenuntersuchungen*. In: Jude, N./Hartig, J. and Klieme, E. (Eds.): *Kompetenzerfassung in pädagogischen Handlungsfeldern*. Bonn.
- Schaeper, H. and Wolter, A. (2008): *Hochschule und Arbeitsmarkt im Bologna-Prozess – Der Stellenwert von „Employability“ und Schlüsselkompetenzen*. *Zeitschrift für Erziehungswissenschaft* 11 (4), 607-625.
- Schomburg, H. and Teichler, U. (2006): *Higher Education and Graduate Employment in Europe*. Dordrecht.
- Statistisches Bundesamt (Ed.) (2007): *Statistik und Wissenschaft. Amtliche Hochschulstatistik und Hochschulrankings*. Wiesbaden.
- Teichler, U. (2002): *Hochschulbildung*. In: Tippelt, R. (Ed.): *Handbuch Bildungsforschung*. Opladen.
- Teichler, U. (2004): *The changing debate on internationalisation of higher education*. *Higher Education* 48, 5-26.
- Teichler, U. (Ed.) (2007a): *Careers of University Graduates*. Dordrecht.
- Teichler, U. (2007b): *Studium und Berufschancen: Was macht den Unterschied aus?* *Beiträge zur Hochschulforschung* 29 (4), 10-31.
- Teichler, U. (2008): *Der Jargon der Nützlichkeit. Zur Employability-Diskussion im Bologna-Prozess*. *Das Hochschulwesen*. 56, 68-79.
- Wissenschaftsrat (2006): *Empfehlungen zum arbeitsmarkt- und demographiegerechten Ausbau des Hochschulsystems*. Köln.
- Wolter, A. (2004): *From State Control to Competition. German Higher Education Transformed*. *The Canadian Journal of Higher Education* XXXIV, 73-104.
- Wolter, A. (2007a): *From the Academic Republic to the Managerial University – The Implementation of New Governance Structures in German Higher Education*. In: University of Tsukuba, Research Center for University Studies (Ed.): *The 3<sup>rd</sup> International Workshop on Reforms of Higher Education in Six Countries/Regions – Commonalities and Differences*. Tokyo.

- Wolter, A. (2007b) Diversifizierung des Weiterbildungsmarktes und Nachfrage nach akademischer Weiterbildung in Deutschland. Österreichische Zeitschrift für Hochschulentwicklung 2 (1), 14-29.
- Zeitschrift für Soziologie der Erziehung und Sozialisation (ZSE) (2008): Der Bologna-Prozess – Potenziale und Folgen für Hochschule und Studium. Vol. 28/No. 4. München.





## 3.6 Adult Education and Lifelong Learning

*Corinna Kleinert and Britta Matthes*

Contact:

Corinna Kleinert  
Institute for Employment Research (IAB)  
Regensburger Str. 104  
90478 Nuremberg  
Germany  
e-mail: [corinna.kleinert\[at\]iab.de](mailto:corinna.kleinert[at]iab.de)

## **Abstract**

Over the last years, political and scientific debates have stressed the growing importance of adult education. Currently important research questions call not only for data sources that collect detailed information on adult education with repeated measurements and in different cohorts, but they should also include data on other life spheres such as education and working histories, partnership and household information, as well as competence development.

In Germany, there are several large-scale datasets containing information on adult education. While general panel studies do not provide a systematic overview of educational activities of adults, studies focusing on adult education are either small-scale or cross-sectional and contain little contextual information. A study that covers information on all educational activities in the life course as well as repeated competence assessment is still missing.

In part, these deficits will be resolved by the large-scale longitudinal studies focused on adults and education that were either recently conducted or are currently being prepared. Thus, we do not call for new data sources on adult education. What is far more important in the next few years is analyzing the data of the new large-scale data sources thoroughly, but also developing new theoretical approaches to adult education.

**Keywords:** adult education, further education, lifelong learning, continuing training, life course, competencies, data access

## **1. The need for analyses in the area of adult education**

Over the last few years, political as well as scientific debates have stressed the growing importance of adult education (Becker and Hecken 2005; European Commission 2000). The significance of this area is largely justified with reference to ongoing globalization, skill-biased technological change, and the development of the knowledge society – changes that have crucial effects on the working lives of the population in (post-)industrial countries. Education is no longer viewed as an asset achieved in youth that remains of constant value during an uninterrupted and stable employment career. Today and in the future, adults must learn continuously to keep up with the flexible requirements of the workplace and to be able to find employment in different and rapidly developing fields.

As a result, there is an urgent political need for knowledge about how to achieve the following goals: How can we enlarge the skill potential among those that have been largely underexploited up until now (e.g., the unemployed or marginally employed, low-skilled, or older workers)? Do these groups have the necessary prerequisites for continuous learning, particularly

in terms of basic skills? How can we ensure that higher-skilled adults continue to learn after completing their initial education? How can they be enabled to flexibly adapt to the changing requirements and new technologies of working life, beyond their employers' immediate needs? How can lifelong learning be organized efficiently in society as a whole so that it reaches all groups of individuals – integrating all the different institutions and organizations involved in adult education? A great deal of empirical research is needed to answer these questions.

In contrast to this evident need, the first national report on education in Germany devoted only sixteen pages to adult education due to the “particularly difficult data situation of adult education” (Konsortium Bildungsberichterstattung 2006: 123). The report merely covered diverging participation rates of subgroups in different types of adult education in Germany, differentiated by educational background, age, gender, occupation, and position (Kuwan et al. 2006). Thus, we know that lifelong learning increases with growing education, occupation, and position, and declines with age (e.g., Pfeiffer and Pohlmeier 1998; Schömann and Becker 1995; Bellmann and Leber 2003; Schiener 2006). Furthermore, men participate more frequently in adult education than women, and natives more frequently than persons with a migration background (e.g., Pfeiffer and Pohlmeier 1998; Becker 2003). In comparison with other nations, we know that the participation rate in adult education in Germany is relatively low (OECD 2006). Finally, the costs of adult education in Germany are mainly borne by firms, the participants themselves, and the Federal Employment Agency (BA, *Bundesagentur für Arbeit*), but figures on financing vary according to different data sources (Beicht et al. 2005).

Apart from these facts, there are many research questions, particularly of a longitudinal nature, that have not yet been answered. This deficit becomes apparent for example when searching for results on *(cumulated) long-term returns of educational activities* in youth as well as in adult life. Since most of the existing data is only cross-sectional, this issue cannot be analyzed. The exceptions to this deficit are the training schemes financed by the BA. In various evaluation studies the success of these programs with regard to subsequent labor market integration is analyzed (e.g., Hujer et al. 2006; Schneider and Uhlenborff 2006). While the returns of (adult) education are mostly understood economically – for example by analyzing income, wages, labor market integration, mobility, or career development – a more pedagogically oriented approach would ask for *learning outcomes*. To answer this question, we need not only precise data on learning activities, but also on the development of individual competencies (for details see Trautwein and Schoon in this publication). Studies targeted at adult persons that combine both topics in a longitudinal design are currently underway (see section 4).

Learning activities are embedded in the adult life course. Up until now we have known little about these *framing mechanisms of adult education* and how they interact with participation, since research on educational pathways and how they are embedded in employment histories and other life domains is still at an early stage (Jacob 2004; Hillmert and Jacob 2004). The decision to participate in further education is connected to specific personal circumstances. Certain factors, such as unemployment, promote participation due to an expected increase in employment chances. Other circumstances may reduce participation due to the time restrictions they impose (e.g., child-bearing) or because of the expectation that such education does not pay off anymore (e.g., in the case of older persons). To explore these research questions, rich data sources in a life course oriented framework are needed.

Another aspect of framing adult education within the life course is the *household*, since it determines the opportunities and restrictions surrounding participation in several respects: first, economic resources and their allocation among household members determine participation; second, the division of labor within households and partnerships has an impact on participation. Thus, the relative position of household members, their educational resources, and their time budgets decide about participation in adult education. These aspects are particularly important for assessing gender differences in lifelong learning. Another household characteristic that is primarily significant for self-learning processes is the learning environment at home. Finally, the household situation not only influences participation in adult education, but even more it affects the *decision-making processes that precede it*. In sociology particularly, there are highly-developed theories and many empirical results related to parental decisions about the educational choices of their children, but far less research is available concerning educational decisions made during adulthood.

Another important research question addresses further education in Germany among adults with a migration background. *Migrants and their descendants* are a group that, in part at least, is urgently in need of education during adulthood since their educational endowment is often inadequate and certificates acquired in their countries of origin are frequently not recognized in Germany.

Finally, many countries are struggling with an aging society; this is especially true in Germany (Fuchs and Dörfler 2005). Thus, it must be ensured that the older population will be equipped for participating in working life longer than it is today by providing access to continuing education. Yet, in contrast to these necessities, we find that older people participate less in continuing learning than younger ones (Kuwan et al. 2006). Thus, it is important that research can identify the *opportunities and barriers to continuing education in older age groups*. To do so, it is necessary

to gain additional knowledge about the returns of educational activities later in life.

Taken together, these research questions call not only for data sources that collect detailed information on adult education with repeated measurements and in different cohorts, but they should also include data on other life spheres, such as education and working histories, partnership and household information, and competence development in different domains.

## 2. A complex field of research

The main challenge that faces data collection in the field of adult education is the complexity of the object of investigation. It is therefore fruitful to distinguish it analytically before describing the relevant data sources. For this purpose we classify adult education according to the form of learning, the learning location and context, and the purpose and contents of learning (Wohn 2007).

Generally, adult education can be defined as “the continuing or resumption of formal, non-formal, and/or informal learning with general or vocational content after completion of initial training” (Expertenkommission Finanzierung Lebenslangen Lernens 2002: 56). *Formal education* is institutionalized and leads to recognized certificates that strongly determine labor market chances in Germany. Therefore most existing data sources are limited to this type of education and the data situation here is well developed. A second – and in quantity and quality more important – type of adult education is *non-formal education*, which includes shorter institutionalized training courses that do not lead to certificates (or to certificates not fully recognized). This is the type of educational activity that is commonly understood when referring to “adult education.” However, data on non-formal education is more difficult to collect: Participation differs individually and problems of recall and identification of these events are common. Even less is known about *intentional informal learning*, learning processes organized by the individuals themselves (e.g., by participating in conferences, reading textbooks, or learning a new computer program). This is particularly true regarding the decisions that lead to these learning processes or their (cumulative) returns. In this context, it is important to mention the limitations of standardized survey research. First, people have difficulty remembering such activities over a longer period of time. Thus, information on non-formal and informal learning can only be collected in a panel design or for a limited retrospective period. Furthermore, survey questionnaires cannot measure *unintentional informal learning* that takes place in the context of other activities – at least not directly. Still, we assume that this form of learning is

very important, particularly when it takes place on the job. Many adults constantly obtain new skills, typically without being aware of it, simply by fulfilling their tasks and responding to the challenges of everyday working life or by performing voluntary activities. Thus, unintentional informal learning can be assessed only indirectly by measuring employment experience, activities and requirements on the job, and social engagement.

While most individuals participate in learning in earlier educational stages within the same predefined institutional contexts, learning processes of adults happen in a multitude of different *learning environments*. Firms are the major providers of adult education in Germany. Thus, certain kinds of information on participation in firm-based training and education cannot be accessed by individual and household surveys alone, but also by firm-level data (see Joachim Wagner in this publication). Other institutions are important providers of adult education as well. Second-chance programs (*Zweiter Bildungsweg*) allow people to complete upper-secondary qualifications and to proceed to tertiary education (evening schools, adult apprenticeships). Upgrade training for employed workers is offered by the Chamber of Industry and Commerce (IHK, *Industrie- und Handelskammer*) or by the Chamber of Crafts and Trades (HWK, *Handwerkskammer*) and allows for the acquisition of additional formal certificates in relatively short courses. A variety of shorter and longer training programs aiming mainly at reintegrating unemployed persons into the labor market are provided by the BA. Adult Learning Centers (*Volkshochschulen*) provide courses in many areas of self-development including languages, art and music, political developments, and information technology. Non-vocational adult education is also provided by a variety of voluntary and non-governmental organizations including religious groups. These courses often target specific sections of the population, such as women or migrants. This list of examples shows how difficult it is to gain a complete overview of the providers of adult education in Germany. This variety also limits the possible information on the institutional contexts of adult education in empirical data.

Finally, adult education covers many *fields*, ranging from basic cognitive competencies to vocational and non-cognitive skills. Adults do not only participate in further education with the objective of vocational training, but also for personal reasons. Researchers, however, are interested mainly in adult education relevant to working life, employability, active participation in society, or coping with everyday life. Whereas formal training undertaken for these purposes can theoretically be distinguished clearly from educational activities taken up for private reasons, such a distinction is not possible for non-formal education. Taking a foreign language course can, for example, be of central importance for the career advancement of one individual, whereas for another person it serves mainly private interests and has no effect on her or his further working life. This problem calls for relatively broad questioning

strategies on the one hand, and for a detailed collection of the contents of further education on the other hand.

This brief overview already suggests that the data situation in the field of adult education may be both confusing and limited. In the next section, we describe and evaluate the most important data sources and their accessibility before presenting new developments in national and international data collection.

### 3. Status quo: Databases and access

In the following, we distinguish between the actors involved in adult education (providers, firms, or individuals) on the one hand, and cross-sectional and longitudinal data on the other hand. Our discussion of datasets is largely focused on German research and includes only selected examples of international and comparative studies.<sup>1</sup>

In Germany, a comprehensive *statistic* on adult education does not exist. Rather, different statistics are found that are only partly compatible (Weishaupt and Fickermann 2001), since they differ in definitions, variables, periods, etc. Official statistics, for instance, the Statistics of the General Education Schools (evening schools), the Statistics of Technical Schools, or the Statistics of Vocational Education provide information on the number of participants and their socio-demographic characteristics within the respective school types. In addition, the manifold providers of adult education produce statistics relating to their own programs (for example, the German Institute for Adult Education (DIE, *Deutsches Institut für Erwachsenenbildung*) generated the Statistics of the Adult Education Program).

More data sources are available for firms, the most important group of adult education providers. Regarding cross-sectional firm-level data, most important to mention are the IW Survey on in-firm further training conducted by the Institute of the German Economy (IW, *Institut der Deutschen Wirtschaft*) and the Continuing Vocational Training Survey (CVTS) con-

---

1 An overview on data sources concerning adult education in other countries is found in the article by Kristen and colleagues (Kristen et al. 2005). Additionally, Statistics Canada (<http://www.statcan.gc.ca>), the US National Center for Education Statistics (<http://nces.ed.gov/>), the Longitudinal Studies Centre (<http://www.iser.essex.ac.uk/survey>) and the Inter-University Consortium for Political and Social Research (<http://www.icpsr.umich.edu/ICPSR/access/index.html>) in the UK, and the Data Archiving and Networked Services in the Netherlands (<http://www.dans.knaw.nl/en/>) offer online information on studies referring to adult education. A web guide made available by the Mannheim Centre for European Social Research (MZES) supports searching for metadata of major European socioeconomic surveys (<http://www.mzes.uni-mannheim.de/projekte/mikrodaten/drafts/index.html>).



ducted by the Federal Institute for Vocational Education and Training (BIBB, *Bundesinstitut für Berufsbildung*). The IW Survey contains data on the provision of workers' educational activities initiated or financed by the firm (such as on-the-job training, reading literature, participation in internal or external seminars, informative meetings, or retraining) (Werner 2006). However, it is hard to gather information on the survey because there is no systematic overview and its data is not yet available. CVTS is a firm survey on the European level containing information on participation rates, hours, costs, and socio-demographic characteristics of the participants. The advantage of CVTS is its voluminous structural information on further education provided and financed by firms, but its value and comparative possibilities are limited, mainly by methodological problems, for example the change of research unit (firm vs. establishment), or the probable higher response rate of firms or establishments that do provide further education.

In contrast to these cross-sectional firm surveys, the IAB Establishment Panel is an annual panel survey of nearly 16,000 German establishments. One of main topics surveyed in this multi-issue study is further education, including information on evaluation of employee demands, provision of internal or external courses, on-the-job training, and participation in self-learning activities. Additionally, data on participant characteristics is available. Since the panel contains a wide range of firm characteristics, it also allows for an analysis of firm-based training in a longitudinal research design. The IAB Establishment Panel is available through the Research Data Center of the Federal Employment Agency at the Institute for Employment Research (IAB, *Institut für Arbeitsmarkt- und Berufsforschung* within the BA, *Bundesagentur für Arbeit*).

Regarding individual data, the most important data source in Germany has been the *Berichtssystem Weiterbildung* up until now, a national, repeated, cross-sectional survey dedicated specifically to further education (Kuwan et al. 2006). Its data – available for the years 1979 and 1988 to 2003 via the data archive at the Leibniz Institute for the Social Sciences (GESIS, Leibniz-Institut für Sozialwissenschaften) – mainly supports analyses of participation in adult learning. About 7,000 respondents have been asked every three years about their participation in a broad range of educational activities and their learning interests. Since 1994, non-formal education has been included, and questions about self-regulated learning have been asked since 2000, containing instruments to record learning environments, learning dispositions, and support by other persons. However, due to its cross-sectional structure, longitudinal analyses on educational careers are not possible.

Another large-scale dataset containing information on adult education activities of individuals is the German Microcensus, a one-percent sample of all German households conducted yearly by the Federal Statistical Office. Between 1970 and 1995, respondents were asked every two years about

further education over the previous two years; since 1996 these questions have been surveyed on an annual basis. The advantage of the Microcensus is the obligation to participate as well as the high number of respondents, so that even results focusing on small subgroups are reliable. Unfortunately, in the Microcensus adult education is restricted to “training, further education and retraining” and there is no information on when exactly the relevant events took place. Furthermore, comparisons between different survey years are limited due to changing instruments and time references. Except for two four-year panel files (1996-1999, 2001-2004) covering 25 percent of the yearly sample, analyses are restricted to cross-sectional designs. The Microcensus data is available via the Research Data Center of the Federal Statistical Office, the Scientific Use Files via the German Microdata Lab at GESIS.

Finally, the *BIBB/IAB Surveys* are a series of large-scale, representative cross-sectional surveys of huge samples of the employed conducted in 1979, 1985/86, 1991/92, and 1998/99. Like all other previous waves, the most recent survey from 2006 (*BIBB-BAuA Survey*) will be available in 2009 via the data archive at GESIS. These surveys gathered rich representative information on qualification profiles and occupational developments, as well as the organizational, technological, and qualification frameworks at the workplace. They also contain limited retrospective data on former educational careers, in particular on initial training. The data on adult education is cross-sectional as well, but has the advantage of capturing formal, non-formal and informal training, as well as activities and requirements of the current job that can be used indirectly as proxies for informal learning activities.

In the field of cross-sectional individual survey data, the situation in Germany can be compared to many other Western countries. Regularly implemented surveys focused on adult education and available for scientific use can be found in the UK with the National Adult Learning Survey,<sup>2</sup> in Finland with the Adult Education Survey,<sup>3</sup> in Sweden’s Staff Training Statistics,<sup>4</sup> or in the US Adult Education Survey.<sup>5</sup> In the future, national surveys in Europe – in Germany the Berichtssystem Weiterbildung – will be replaced by a common data source, the European Adult Education Survey (AES). AES was carried out for the first time in 2007 on a voluntary basis in over twenty European countries and provides information about adult participation in formal, non-formal, and informal training. The first round of obligatory data collection will be in 2011 (e.g., Gnahn et al. 2008; Rosenblatt

---

2 <http://www.statistics.gov.uk/StatBase/Source.asp?vlnk=1329&More=Y>

3 [www.stat.fi/meta/til/aku\\_en.html](http://www.stat.fi/meta/til/aku_en.html)

4 [http://www.scb.se/Pages/Product\\_\\_\\_\\_9001.aspx](http://www.scb.se/Pages/Product____9001.aspx)

5 <http://nces.ed.gov/nhes/>

and Bilger 2008). The German AES data is available at the data archive at GESIS.

In the US, the long tradition of student assessment has also led to a comprehensive literacy assessment study focused on the adult population, the National Assessment of Adult Literacy (NAAL),<sup>6</sup> which was carried out in 1992 and 2003. This survey also set the groundwork for international studies on adult skills, learning, and competencies, including the International Adult Literacy Survey (IALS) from the mid 1990s with twenty-two participating countries, and the six-country Adult Literacy and Lifeskills Survey (ALL)<sup>7</sup> carried out in 2003. These studies combine questionnaire data on educational qualifications and different forms of learning with assessments of basic cognitive domains, such as reading literacy or numeracy. Germany participated in IALS, but not in ALL.

Longitudinal datasets on adult education from an individual perspective are available as well. Rich data on educational and employment careers can be found in the German Life History Study (GLHS) of the Max Planck Institute for Human Development, which collected retrospective data on educational, employment, and family histories of several birth cohorts (from the 1920s to 1971). A Scientific Use File is available via the data archive at GESIS or by contacting the Center for Research on Inequality and the Life Course (CIQLE) at Yale University. However, it is well known that the recall of continuing education, in particular of short or relatively minor courses, is restricted. Thus, the extent of non-formal educational activities is underestimated in this survey and probably systematically selective.

More respondents are interviewed in the German Socio-Economic Panel (SOEP, *Sozio-oekonomisches Panel*), a large general household panel survey that has been carried out every year since 1984. The data is made available to researchers by the Research Data Center of the SOEP. This survey focuses on economic issues and employment careers by combining retrospective information with panel data. Further education was the main topic in 1989, 1993, 2000, and 2004. These panel waves cover information on participation in adult education, the number of courses, their extent and duration, goals, providers, costs and financing, with additional questions on the general motivation to participate in adult education. However, the instruments were mainly focused on formal and non-formal training (Pischke 2001) and further education was not linked to the employment history or the employer (Kuckulenz 2007).

Another longitudinal dataset that can be used for analyzing a particular type of adult education – programs provided by the BA – is the Integrated Employment Biographies Sample (IEBS, *Stichprobe der Integrierten Erwerbsbiografien des IAB*). This process-produced dataset contains obser-

---

6 <http://nces.ed.gov/naal/>

7 <http://nces.ed.gov/surveys/all/>

vations on employment, unemployment benefits, job search, and participation in active labor market programs on a daily basis, combining records from four data sources: the IAB employment history, the IAB benefit recipient history, the participants-in-measures data, and data on job search originating from the applicant pool database. Thus, the IEBS enables detailed longitudinal analyses of the participation in measures of active labor market policy. This dataset is available through the Research Data Center of the Federal Employment Agency at the Institute for Employment Research (IAB, *Institut für Arbeitsmarkt- und Berufsforschung* within the BA, *Bundesagentur für Arbeit*).

In the field of longitudinal data on individuals, Anglo-American countries, having launched birth cohort panel studies focused on educational pathways already decades ago, play a leading role today. In the UK, these panels started with newborns (the National Child Development Study, or NCDS, with birth cohort 1958 and the British Cohort Study, or BCS, with cohort 1970<sup>8</sup>), while US studies began primarily with high school students (the National Longitudinal Study of the High School Class of 1972, or NLS-72; the High School and Beyond Longitudinal Study of 1980, or HS&B; and the National Education Longitudinal Study of 1988, or NELS:88<sup>9</sup>). Both approaches have certain disadvantages. The UK panel studies were followed up only in long intervals during adult life. The US surveys concentrate on transitions from training and higher education into employment and usually stop following up their respondents after their mid-twenties. Thus, these data sources are suitable only to a limited extent for analyzing questions about adult education.

To sum up, there are several large-scale datasets containing information on adult education in Germany. Since the publication of the initial expert report and recommendations on the improvement of the information infrastructure in 2001, the possibilities for data access have improved considerably. However, a representative longitudinal study with the main focus on educational issues, such as the birth cohort studies in Anglo-American countries, is still missing. Most large-scale panel studies in Germany still have a broader focus and thus do not provide a systematic overview of educational activities of adults in all panel waves. Studies focusing on adult education are either small-scale or cross-sectional and contain little context information. Moreover, most data sources do not cover all sources of educational activities, and thus do not provide a comprehensive view of educational histories over the life span. Finally, the field of adult education also includes the aspect of lifelong learning, at least from an educational science perspective. This view calls for instruments measuring competence attainment and development. A study that covers both – information on educa-

---

8 <http://www.cls.ioe.ac.uk/>

9 <http://nces.ed.gov/surveys/>

tional activities as well as the repeated assessment of competencies – is still lacking.

#### 4. Future developments

These deficits will be partially resolved by large-scale surveys that have either been recently conducted or are currently being prepared: the IAB study Changing Conditions of Working and Learning (ALWA, *Arbeiten und Lernen im Wandel*), the adult stage of the National Educational Panel Study (NEPS), and the international survey PIAAC.

ALWA was designed to study relationships between formal education, basic cognitive skills, and the working life of adults from a longitudinal perspective (Kleinert et al. 2008). It focuses on recording detailed education and employment biographies of the respondents and on testing their literacy and numeracy skills. The design combines these two components in the form of computer-assisted telephone interviews and paper-and-pencil personal interviews. The target group of the survey is the German population, age 18 to 50. In the 2007-2008 survey, 10,000 persons, chosen on the basis of a random sample from the Resident Registration Offices, were questioned by telephone, and a subsample of 4,000 persons participated in the skills tests. In the CATI questionnaires, all formal educational activities over the whole life course were surveyed. Questions on non-formal education were integrated into the modules on employment, unemployment, and other events to ensure better recall. In addition, data on informal learning activities were collected for the last two years before the interview. Due to its complex structure, the dataset will not be made publicly available (via the Research Data Center of the BA at the IAB) until mid-2010.

From 2009 on, the ALWA participants will be followed up in the context of the adult stage of the National Educational Panel Study (NEPS), in a panel design with yearly intervals (for a description of the complete study, see Blossfeld in this publication). Additionally, the sample will be extended: the study will cover all adults (including migrants) of working age older than twenty-two years, regardless of employment status. In contrast to the ALWA study, the NEPS adult stage is more strongly focused on adult education and lifelong learning. Thus, it is planned to design, test, and implement more detailed instruments covering non-formal and informal learning activities for the retrospective period between panel waves, and to supplement them by indirect measures of informal learning such as job tasks and requirements, and volunteering. One of the main goals of NEPS is to make its data publicly available quickly after the data is gathered.

A second new development in the field of adult training was inspired by the realization that we need to know more about the providers of adult education in order to learn about training decisions and learning processes. Considering the multitude of actors in the field, this would be a difficult goal to achieve in the case of adult education in general; however, it is a more reasonable goal for firm-based training and education. One approach to this is combining individual and firm-level data, a method that is currently implemented in projects linking individual survey data and administrative data (for a detailed discussion, see Schnell in this publication). For instance, ALWA and NEPS will use record-linkage routines to enrich respondents' data on employment periods with establishment information from administrative data.

Another approach to data linkage was undertaken in the project 'Further Training as a Part of Lifelong Learning (WeLL)'. This project of RWI Essen, IAB, infas, and DIE aims at analyzing the joint training decisions of employers and their employees (Bender et al. 2008a; b). First, an employer survey was conducted in 2007, followed by a panel survey of employees in the respective firms. Both surveys focus on the collection of training information together with a variety of employee and employer background characteristics. Moreover, administrative longitudinal employee data can be linked with these data sources. In 2010, the project will provide its data via the Research Data Center of the BA at the IAB.

Finally, a large-scale international OECD survey on adult education is currently being prepared. The Programme for the International Assessment of Adult Competencies (PIAAC) will assess the level and distribution of adult skills by focusing on key cognitive and workplace skills across countries. PIAAC will also gather information on the antecedents and outcomes of skills, as well as on the use of information technology and literacy and numeracy practices in general. Its data will allow researchers to investigate links between key cognitive skills and a range of demographic variables, economic and other outcomes, as well as the use of skills in the workplace and other settings. The survey will be administered in 2011 and its results are scheduled to be released in early 2013.

## **5. Two final recommendations**

The comparison between two important data sources on adult education in Germany results in astonishing disparities, even in terms of basic information. For example, according to the data of the Berichtssystem Weiterbildung IX, 41 percent of the adult population in Germany participated in further education in 2003, while the Microcensus reported only 13 percent

(Wohn 2007). The main reason for this significant discrepancy seems to be the highly different instruments of the two surveys. This problem arises not only around these particular studies. Most other surveys use specific, non-comparable instruments as well. Often, they are constructed ad hoc and not sufficiently tested. In part, this problem is simply a reflection of the complexity of adult education and its 'resistance' to standardized survey research. Thus, an important challenge to be met over the next few years is to develop standardized, valid, and reliable instruments representing the entire range of educational activities in adulthood, at least as far as they are undertaken intentionally and can be recalled. To a certain extent, these development tasks are central to the above mentioned new studies – above all to the National Educational Panel Study (NEPS). Moreover, the call for standardized instruments has also an international dimension. To date, the results of most German studies cannot be compared internationally, since instruments and item batteries differ considerably by country. This is not only a problem of poor international coordination, but also of different national meanings of adult education, educational cultures, and institutional conditions. Here, new international studies such as PIAAC could take a lead in helping to integrate instruments.

Since we are now (it is hoped) taking a step forward with tackling these problems by way of the new surveys mentioned above, we are not issuing a call for new data sources on adult education. What is far more important in the next years, in our view, is testing these new large-scale data sources, analyzing the data thoroughly, but also developing new theoretical approaches to adult education. It is from these areas that we will find the greatest challenges in the upcoming years. More researchers from diverse fields – including sociology, economics, psychology, and educational science – should work with innovative theoretical approaches and state-of-the-art empirical methods on the existing and new data to generate more knowledge about adult education and to explore its development and its relationship with structural changes in the labor market and the life course. This calls for a strong initiative in the training and promotion of young empirical researchers in these fields.

## References:

- Becker, R. (2003): Educational Expansion and Persistent Inequalities of Education: Utilizing Subjective Expected Utility Theory to Explain Increasing Participation Rates in Upper Secondary School in the Federal Republic of Germany. *European Sociological Review* 19 (1), 1-24.
- Becker, R. and Hecken, A. (2005): Berufliche Weiterbildung – arbeitsmarktsoziologische Perspektiven und empirische Befunde. In: Abraham, M. and Hinz, Th. (Eds.): *Arbeitsmarktsoziologie: Probleme, Theorien, empirische Befunde*. Wiesbaden, 133-168.
- Beicht, U./Berger, K. and Moraal, D. (2005): Aufwendungen für berufliche Weiterbildung in Deutschland. *Sozialer Fortschritt* 54 (10/11), 256-266.
- Bellmann, L. and Leber, U. (2003): Betriebliche Weiterbildung. Denn wer da hat, dem wird gegeben. IAB-Materialien, 1/2003. Nürnberg.
- Bender, S./Fertig, M./Görlitz, K./Huber, M./Hummelsheim, S./Knerr, P./Schmucker, A. and Schröder, H. (2008a): WeLL – Berufliche Weiterbildung als Bestandteil Lebenslangen Lernens. Projektbericht. FDZ Methodenreport, 5/2008. Nürnberg.
- Bender, S./Fertig, M./Görlitz, K./Huber, M. and Schmucker, A. (2008b): WeLL – Unique Linked Employer-Employee Data on Further Training in Germany. *Ruhr Economic Paper* 67. Essen.
- European Commission (2000): A Memorandum on Lifelong Learning. Brussels.
- Expertenkommission Finanzierung Lebenslangen Lernens (2002): Auf dem Weg zur Finanzierung Lebenslangen Lernens. Zwischenbericht. Schriftenreihe der Expertenkommission Finanzierung Lebenslangen Lernens, Vol. 1. Bielefeld.
- Fuchs, J. and Dörfler, K. (2005): Projektion des Erwerbspersonenpotenzials bis 2050: Annahmen und Datengrundlage. IAB Forschungsbericht, 25/2005. Nürnberg.
- Gnahn, D./Kuwon, H. and Seidel, S. (2008): Weiterbildungsverhalten in Deutschland. Band 2. Berichtskonzepte auf dem Prüfstand. Bielefeld.
- Hillmert, S. and Jacob, M. (2004): Multiple Episodes: Training Careers in a Learning Society. Universität Bamberg Globalife Working Paper 64. Bamberg.
- Huier, R./Thomsen, S.L. and Zeiss, Ch. (2006): The effects of vocational training programmes on the duration of unemployment in Eastern Germany. *Allgemeines Statistisches Archiv* 90, 299-321.
- Jacob, M. (2004): Mehrfachausbildungen in Deutschland. Karriere, Collage, Kompensation? Wiesbaden.
- Kleinert, C./Matthes, B. and Jacob, M. (2008): Die Befragung „Arbeiten und Lernen im Wandel“. Theoretischer Hintergrund und Konzeption. IAB-Forschungsbericht, 5/2008. Nürnberg.
- Konsortium Bildungsberichterstattung (Ed.) (2006): Bildung in Deutschland. Ein indikatorengestützter Bericht mit einer Analyse zu Bildung und Migration. Bielefeld.
- Kristen, C./Römmel, A./Müller, W. and Kalter, F. (2005): Längsschnittstudien für die Bildungsberichterstattung – Beispiele aus Europa und Nordamerika. Berlin.
- Kuckulenz, A. (2007): Studies on Continuing Vocational Training in Germany. Mannheim.



- Kuwan, H./Bilger, F./Gnahn, D. and Seidel, S. (2006): Berichtssystem Weiterbildung IX: Integrierter Gesamtbericht zur Weiterbildungssituation in Deutschland. Berlin.
- OECD (2006): Education at a Glance. Paris.
- Pfeiffer, F. and Pohlmeier, W. (Eds.) (1998): Qualifikation, Weiterbildung und Arbeitsmarkterfolg. Baden-Baden.
- Pischke, J.-St. (2001): Continuous Training in Germany. *Journal of Population Economics* 14 (3), 523-548.
- Rosenblatt, B. v. and Bilger, F. (2008): Weiterbildungsverhalten in Deutschland. Band 1. Berichtssystem Weiterbildung und Adult Education Survey 2007. Bielefeld.
- Schiener, J. (2006): Bildungserträge in der Erwerbsgesellschaft. Wiesbaden.
- Schneider, H. and Uhlendorff, A. (2006): Die Wirkung der Hartz-Reform im Bereich der beruflichen Weiterbildung. *Zeitschrift für Arbeitsmarktforschung* 39 (3/4), 477-490.
- Schömann, K. and Becker, R. (1995): Participation in Further Education over the Life Course. A Longitudinal Study of Three Birth Cohorts in the Federal Republic of Germany. *European Sociological Review* 11 (2), 187-208.
- Weishaupt, H. and Fickermann, D. (2001): Informationelle Infrastruktur im Bereich Bildung und Kultur (Expertise für die Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik. In: Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik (KVI) (Ed.): Wege zu einer besseren informationellen Infrastruktur. Baden-Baden.
- Werner, D. (2006): Trends und Kosten der betrieblichen Weiterbildung – Ergebnisse der IW-Weiterbildungserhebung 2005. *IW-Trends – Vierteljahresschrift zur empirischen Wirtschaftsforschung* 33 (1).
- Wohn, K. (2007): Effizienz von Weiterbildungsmessung. RatSWD Working Paper No. 19.



### 3.7 Research, Science, Development

*Stefan Hornbostel*

Contact:

Stefan Hornbostel  
Humboldt-Universität Berlin  
Institut für Sozialwissenschaften  
10099 Berlin

iFQ - Institut für Forschungsinformation und Qualitätssicherung  
Godesberger Allee 90  
53175 Bonn  
Germany

e-mail: [hornbostel\[at\]forschungsinfo.de](mailto:hornbostel[at]forschungsinfo.de)

## Abstract

As in other societal realms, in research, science, and development, governments have been increasingly placed under pressure to legitimize their actions. Accordingly, it is only natural that governments wish to base future activities on well informed and empirically grounded decisions. As a result, demand for performance measures, benchmarking, comparative analysis, and “foresight studies” has grown significantly. To meet this demand, rankings, ratings, and evaluations have been supposedly introduced to, on the one hand, produce transparency and, on the other hand, act as stimuli to improve performance. To date, however, central questions relating to the underlying methodologies and indicators used in evaluation measures have been left unanswered. These questions concern not only the availability and appropriateness of the data, indicator construction, and methodologies used, but also how to approach effects caused by disciplinary, sectoral, regional or national differences. Furthermore, questions regarding the intended and unintended effects of the evaluation instruments used have also been left unresolved. This article will describe and discuss these issues in greater detail. In Germany, infrastructural deficiencies, such as the fragmentation of research groups, have prevented open research questions from being addressed. Within this context, the two most important tasks identified are the development of a decentralized data collection system that would enable standard definitions, and the development of competitive research infrastructure.

Keywords: science indicators, R&D, research funding, governance

## 1. Research questions

On account of Germany’s federal structure, its research and innovation landscape is both highly diversified and differentiated. Research is conducted in state and non-state institutions, institutes of higher education, non-university research institutions as well as in industry, which alone constitutes two-thirds of invested research funds.

On an international level, interest in assessments and comparative analyses of higher education and research systems has increased significantly since the 1970s. This growth in interest can be attributed to two closely-related factors. First, the development of a knowledge-based society whose interests are increasingly dependent on research and technology has inevitably contributed to the growing trend to evaluate higher education and research. Second, this development has been driven further by the fact that the quality of this research and technological progress in turn relies on the continual search for ideal conditions in which well-qualified junior scientists, innovation, and top-class research can be fostered. Additionally, in the name of legitimacy and planning, governments are progressively demanding more

performance measures, international comparisons, and “foresight studies.” In the course of the 1980s and 1990s, output-orientated research funding gradually increased and, in many European states, the competitive orientation of the academic system grew. Rankings, ratings, evaluations, and formula-based allocation schemes were deemed to provide the necessary transparency, problem diagnoses, and performance-raising stimuli.

Michelson (2006) describes the trend in research assessment in the US as follows:

“First, the standardization and harmonization of performance assessment methodologies has begun to spread across various federal R&D funding agencies. [...] Second, there has clearly been a turn toward employing quantitative methodologies as a major part of performance assessment initiatives. [...] Third, the growing use of quantitative bibliometric indicators is also being paired with a renewed focus on utilizing qualitative indicators in an effort to create more appropriate hybrid methodologies that can capture a wider range of variables related to a program’s performance.”

These three developments can also be observed in Europe.

When conducting analyses of the academic system, results can be systematically contrasted with relevant data or, for example, governance instruments, depending on the purpose. Assessment subjects can range from the academic performance of individuals, of organizations, of institutions (workgroups, institutes of higher education), or of branches of research (research fields, disciplines), through to national academic systems. Essentially, for such analyses, indicators and peer-review procedures are most often employed, as well as a combination of the two (this is known as an informed peer review). Because the use of data and assessment procedures in the academic world is as varied as the actors who conduct or commission them, only a few applications will be mentioned as examples. One such example is the German Council of Science and Humanities (WR, *Wissenschaftsrat*), which publishes nationwide research ratings for selected subjects. These ratings are based on series of output data about assessed research units, which are evaluated by peers according to a uniform scale and various criteria (WR 2008).

However, in contrast to the British Research Assessment Exercise, which utilizes a similar methodological structure, funds in Germany are not allocated according to ratings. Non-university research institutes, such as the Max Planck Society or Leibniz Association, conduct regular assessments of their member institutes and have their own departments for carrying out this task.<sup>1</sup> In such cases, typical indicators for measuring performance, which includes publications, third-party funding, patents, and services, are drawn upon and used to make decisions concerning the allocation of further funds to the establishments. Research funding bodies, such as the German Research

---

1 Max Planck Society: <http://www.mpg.de/ueberDieGesellschaft/profil/evaluation/index.html>,  
Leibniz Assoziation: <http://www.wgl.de/?nid=veva&nidap=&print=0>.

Foundation (DFG, *Deutsche Forschungsgemeinschaft*), the Alexander von Humboldt-Foundation or the European Research Council, regularly evaluate the outcomes of their funding programs on a selective basis.

Additionally, such institutions have established monitoring systems, which regularly provide data concerning the performance of particular programs. The *Länder* and university faculties, for instance, make use of performance-based funding allocations to distribute part of their budget according to positive and negative performance indicators. Such systems are based exclusively on quantitative performance indicators relating to research, and in some cases teaching. Some federal states even have their own assessment centers, such as Lower Saxony's Scientific Commission (WKN, *Wissenschaftliche Kommission Niedersachsen*) or Baden-Württemberg's Evaluation Agency (EVALAG, *Evaluationsagentur*).<sup>2</sup> These assessment centers conduct regular or special-purpose assessments of academic establishments. Institutes of higher education also develop their own evaluation and reporting procedures to collect and disseminate information about performance in teaching and research. They link these assessments to target agreements, which are established by management and the corresponding faculties or institutes.

Until now in Germany, the use of performance indicators has not played a significant role in pay negotiations. But with the introduction of elements of performance-related pay, typical research indicators will play an increasingly important role in this area of individual agreements. In some disciplines already, but especially in the life sciences, the use of specific indicators is being used informally. Examples of this include the application of Journal Impact Factors for measuring publication activity or the Hirsch Index for quantifying individual research performance within the framework of employment and appointment negotiations (Jaeger 2006; Vahl 2008).

With the significant increase in quantitative indicators and the availability of complex indicators, expectations relating to data quality and knowledge of their governing factors have grown considerably, even among non-specialists. The error tolerance when small units are analyzed is drastically lower than when larger units are taken into account. This is also true of the use of indicators (which, it should be noted, are often not assessed) that are frequently used for a different purpose from that which was originally intended. The Journal Impact Factor, for example, was developed to characterize academic journals but is generally usually used to provide an indication of the quality of individual publications. It is difficult to assess the impact of bias effects, especially when small units are being compared, because there are very few foolproof error theories. Moreover, selected procedures and indicators cause learning effects among the concerned aca-

---

2 Wissenschaftliche Kommission Niedersachsen: <http://www.wk.niedersachsen.de>,  
Evaluationsagentur Baden-Württemberg: <http://www.evalag.de>.

demics. Behavior, which is directly geared towards “indicator polishing,” can, although not always, bring about unwanted effects (Moed et al. 2005).

### *1.1 Indicators*

The call to develop appropriate indicators for measuring performance in research and development, as well as measures of potential indicators, was formulated within this context and still applies to this day. Central questions in this area are:

- What mechanisms can be used to measure academic performance? Apart from survey techniques (reputation surveys, Delphi surveys, etc.) and the analysis of funding data, bibliometric mechanisms have been developed as a method of measuring academic performance. Additionally, peer reviews of performance and, in particular, techniques used in patent data analysis have also become increasingly used.
- How can national, disciplinary, and sub-disciplinary specifics be taken into account when using publication and citation analyses as indicators? Publication and citation behavior, the intensity of third-party funding, or patenting strategies differ considerably according to discipline. A clear implication of this is that standardization mechanisms are needed for comparative analyses and descriptions. This point is also highlighted by the use of national languages in academic publications in the larger non-English speaking realm, as this equally makes comparison difficult. Publications in a national language necessarily reach a smaller audience and thus have fewer chances of being cited; a strong argument for developing appropriate indicators. Notably, although bibliometric mechanisms are mainly applied to the life and natural sciences, they are becoming increasingly used in the humanities and social sciences across Europe (Hornbostel 2008a). This trend allows for some degree of comparison, however, this single shared feature is relatively insignificant when compared to the remaining differences.
- How can research performance be assessed in the applied disciplines? Classic bibliometrics has a limited function in these areas and is often substituted by analyses of patenting activities. Here, too, exists a series of problems related to content and methodology. Questions that researchers (Butler 2006; Butler and Visser 2006) are currently working on within this field include: to what extent do the most-used triad patent data relate to the income from license agreements? What patenting strategies are used in



which fields? How are patents related to the academic literature? To what extent do these indicators signal the existence of innovation processes?

- Can standard and internationally-applicable definitions of input, process, and output values be developed? Even within a national context, it is difficult to compare uniform compilations of input values (monetary values, personnel, etc.) due to the heterogeneity of research systems. This is undoubtedly also true of the international arena. In the 1970s, the OECD started standardizing the variables used for evaluating R&D. However, many problems remain unresolved – especially when institutions with different legal and organizational structures are compared.
- International cooperation has become a very important political issue over the course of the past 30 years. Indeed, foreign policy has expanded to include “academic foreign policy.” In light of the growing trend towards specialization, the importance of international academic cooperation will most likely continue. The conditions for successful international cooperation, the consequences of such cooperation, and questions on the methodology of measuring the intensity and impact of international academic cooperation, are some of the chief current questions being posed by researchers and government alike. These questions are addressed mostly within the realm of the aforementioned indicators (co-authorship analyses, international patent announcements, citation networks, CV analyses, mobility analyses) (Schmoch et al. 2006).
- R&D expenditure is evaluated within the framework of official statistics and treated as far as possible according to international standards as specified by the OECD. In Germany, agreements have been met between the WR, the Federal Statistical Office, the Conference of the Ministers of Education and Cultural Affairs (KMK, *Kultusministerkonferenz*), and the Federal Ministry of Education and Research (BMBF, *Bundesministerium für Bildung und Forschung*) to compile data about academic staff (Hetmeier 1998). Questions regarding qualifications and subject expertise cannot, however, be answered with currently available data.
- Third-party funding is harder to assess. Competitive third-party funding that is granted after expert consultant approval is an important research indicator. This is because the approved funding is registered by the recipients and the pertinent funding bodies are also in receipt of the relevant data (Hornbostel 2001; Hornbostel and Heise 2006). After considerable teething problems, the situation concerning third-party funding has improved considerably. Nonetheless, it remains somewhat problematic, especially with regard to European funding (e.g., the Framework Pro-

gramme). In particular, the blurred cut-off line between funds for basic research and those for development or contract research remains controversial. It is argued that the use of third-party funds can lead to considerable bias in disciplines or sub-disciplines that are often only in comparatively limited need for third-party funding. The interpretation of third-party funding indicators also creates problems because only the assessor's evaluation of *quality* is important during the approval process. The actual *quantity* of funds, which can amount to significant investments in a given project's research infrastructure, is often of comparably lesser importance.

- Data about junior scientists, especially the number of PhD candidates, are often used as research indicators. Doctoral candidates often find themselves on the border between teaching and research systems. The Bologna Process regards the doctorate as the third cycle within the academic training process. Unfortunately, apart from the number of completed doctorates, there exist very little data about the quality of academic training and the selection process. Equally, there is very little information available about the career paths of doctoral students. The increasingly used criteria of the number of doctoral students for allocating funds is, therefore, purely quantitative and does not take quality into account. This needs to be addressed urgently (Berghoff et al. 2006; Hornbostel 2008b).
- In the field of innovation research, the central question regarding performance measures revolves less around typical performance measures than it does around the identification of scientific “breakthroughs” and their possible application in products and services. Apart from the issue of how such “breakthroughs” can be recognized from an early stage, there is the related question of what conditions are needed to enable a rapid transfer of knowledge about essential research questions to other social sectors.

R&D data are not sufficient enough to address the aforementioned issues, especially as, for historical reasons, their compilation has very much been geared towards industry. Correspondingly, this makes it difficult to chronicle knowledge-based innovations in the service sector.

### *1.2 Effects analysis / governance*

A second set of questions arises with respect to the topic of appropriate governance structures: which structures best generate conditions ideal for innovative and efficient research? Ideally, these conditions should enable a knowledge transfer between research and other social sectors, as well as

establish linkages between economic growth and the breadth and type of R&D investment. However, the heterogeneity of research and funding systems only allows for analyses that provide limited information because of the lack of compatible data. This problem is exacerbated by several unresolved problems concerning indicators, especially in comparative analyses at an international level.

Over the past 15 years, the governance structures of the higher education system in Germany have changed dramatically. There has long been a shift in research funding due to an increase in third-party funding and simultaneous decrease in access to standard basic equipment. This trend has been augmented by the growing competition among institutes of higher education and within institutes themselves for basic equipment, which, increasingly, is allocated according to performance (Jansen 2007). At the same time, institutes of higher education have gradually gained more and more autonomy whilst having been compelled to develop stronger strategy and management skills. The driving forces behind these developments include the spawn of higher education representative bodies, internal organizations, and management structures, as well as pressure for the creation of a competitive profile in research and teaching. Other contributing factors include the distribution of expertise between the German Federal Government and the *Länder*, the “European research realm,” and the general statutory framework. Additionally, in some cases even the statutory position of institutes of higher education themselves have contributed to the simultaneous growth in autonomy and management strategies.

In light of the aforementioned, information about academic performance has gained greater significance in numerous aspects – as comparative data for stakeholders, as an internal monitoring system, as an instrument of accountability for financiers, and as a component of governance systems (ESF 2008). This is true not only of institutes of higher education, but of all actors in the academic system. Until now, however, the necessary data have been compiled, if at all, *in situ* and according to contrasting standards. Similarly, performance indicators have also been defined in different ways. Technical systems have not been developed with interoperability in mind. This means that while data are often compiled several times, they do not necessarily exist in formats that easily enable their exchange.

### *1.3. Data compilation*

At an early stage, the increased significance of R&D triggered attempts to compile data about input and output variables on a regular basis. The first international “Science and Engineering Indicators Report” was published in the US in 1973 by the National Science Board of the National Science Foundation. The OECD followed this up in the 1980s and has since regularly

published the “OECD Science, Technology and Industry Scoreboard” and the “OECD Science, Technology and Industry Outlook” on an alternating basis. Each publication gives an overview of the trends in science, technology, and innovation policy, all of which are backed by data. Similarly, Eurostat has been compiling data since the beginning of the 1980s with its Science, Technology and Innovation in Europe series. In Germany, the “Report of the Federal Government on Research” (*Bundesbericht Forschung*) publishes information about R&D activity. Regular compilation of data also takes place within Germany at an institutional level or within the framework of research funding. One such example is given by the rankings created by the German Research Foundation (DFG, *Deutsche Forschungsgemeinschaft*) (National Science Board 2006a, 2006b; OECD 2008a, 2008b; BMBF 2008; Europäische Kommission 2008; Statistisches Bundesamt 2008; DFG 2006).

## 2. Status quo: Databases and access

Germany’s “Report of the Federal Government on Research” records the growing demand for contemporary data about the development of investments in research. However, as yet, data about R&D investment tend to be published after considerable delay because data from the federal government, the states, and industry have to be combined. The data are not appropriate for an outcome-oriented analysis. This is due to the fact that, apart from the official data, although there is a wealth of data about rankings, ratings, and evaluations that are compiled more or less regularly, they are limited or not accessible and very different in terms of quality (Hornbostel 2007; 2006). In this regard, the WR is exemplary because it makes its ratings accessible in a format suitable for scientific use (WR 2008).

Publication and citation data are accessible thanks to two large commercial databases (Web of Science and Scopus) and an abundance of specialized subject databases. They do not, however, usually enable citation analysis. Recently, Google Scholar and researchable open access repositories have started providing publication and citation analyses. Many of these databases offer a series of bibliometric codes. But these impressive masses of data hide a series of problems. In Germany, for example, the few small database workgroups which have emerged are ill-equipped in terms of staff and, from a technical point of view, cannot afford to accumulate expertise over the long-term. In other European countries, however, over the past 20 years some extremely well-performing institutes have been established. These institutes, which have in-house databases and the capacity to develop specific indicators, empirically monitor the academic system and engage in

infrequent bibliometric analyses. In light of this, the BMBF is promoting a consortium<sup>3</sup> of German establishments, which are intended to close the gap between German database workgroups and other European institutes by creating a “bibliometric expertise center.”

Research about patent data can be conducted with the German Patent Information System<sup>4</sup> (DEPATIS, *Deutsches Patentinformationssystem*) of the German Patent and Trade Mark Office. The European Patent Office Worldwide Patent Statistic Database (also known as EPO PATSTAT<sup>5</sup>), however, on account of it having been specifically developed for use by governmental/intergovernmental organizations and academic institutions, is more appropriate. Distribution of this database is, however, restricted, and commercial use is not foreseen.

Substantial data compilation about academic performance and related staff and material inputs requires a combination of heterogeneous information from different sources. Sources of information include academics (self-input), institutes of higher education, third-party funding bodies, and bibliometric and patent databases, among others. A prime example of how output data is collected at institutes of higher education and combined with other data is given by the Norwegian research information system, entitled Frida.<sup>6</sup> Since 2004, Frida has been used as a quality-controlled author-based register of research publications and other types of research outputs. The catalyst for Frida’s emergence was the new outcome-based financing system for Norwegian universities and colleges. Norwegian institutions must now document their actions in order to receive a full share of government funding. Frida is associated with the Norwegian Open Research Archives (NORA),<sup>7</sup> which were launched at the same time. The broader objective of these projects is to develop a central Open Archive Initiative harvesting service. It is intended that this service will be open to all Norwegian research institutions that have both online material in full text and metadata in harvestable format.

In contrast to Norway, Germany lacks such a coordinating body for collecting data. In Germany, not only are definitions of data very different, but technical systems have been developed on a decentralized basis and not generally geared towards an exchange of information. It would seem almost inevitable that chaos should ensue.

---

3 Consisting of the Institute for Research Information and Quality Assurance (iFQ, *Institut für Forschungsinformation und Qualitätssicherung*), the Fraunhofer Institute for Systems and Innovation Research (ISI, *Institut für System- und Innovationsforschung*), IWT University Bielefeld, and the Leibniz Institute for Information Infrastructure (FIZ Karlsruhe, *Leibniz-Institut für Informationsinfrastruktur*).

4 [www.depatist.net/](http://www.depatist.net/).

5 <http://www.epo.org/patents/patent-information.html>.

6 <https://wo.uio.no/as/WebObjects/frida.woa/wo/0.0.27.2>.

7 <http://www.ub.uio.no/nora/noaister/search.html?siteLanguage=eng>.

### 3. Conclusions and recommendations

Overall, data compilation about science and research in Europe is far from sophisticated and outcome-oriented. Further, it lacks in comparability, as data are by no means standardized, despite early attempts at standardization by the Frascati Manual (1963). In Germany, like in other European countries, interest in observing, analyzing and evaluating the academic system has increased substantially and this trend will presumably continue. The reasons lie less in an academic interest than in the consequences of higher education and research reforms, which have brought about some serious changes to governance mechanisms. Knowledge of both structures and of the effects of measures undertaken has a significant role to play across the board. Fast-growing competition worldwide, at an academic and technological level, especially from emerging nations, is also increasing the political pressure to act. The competition can already be perceived in the massive shifts in the worldwide distribution of publications, citations, and patents towards emerging countries.

Alongside qualitative analyses and peer review-based expert opinions, quantitative procedures in the compilation, analysis, and evaluation of research data are gaining in importance. There are several reasons for this, which range from an already perceptible overuse of peer reviews to the need for methodological, controlled comparisons and unanimous indicators. Another reason is that certain questions are deemed no longer answerable from the perspective of individual experts.

Overall, the status quo in Germany, in terms of the coordinated collection of data pertaining to the academic system, the training of experts for processing and evaluating this data, and the quality of the data itself, is deplorable. Data about certain important areas simply do not exist, the comparability of existing data is often limited, and in the field of bibliometric analyses Germany risks falling behind. The two most important tasks, therefore, consist, on the one hand, in developing a decentralized data collection system (CRIS, *Current Research Information System*)<sup>8</sup> that will enable standard definitions to be developed and for centrally-compiled data to be interoperable, as well as, on the other hand, developing a competitive research infrastructure.

---

8 <http://www.eurocris.org/>.

## References:

- Berghoff, S./Federkeil, G./Giebisch, P./Hachmeister, C.-D./Hennings, M. and Müller-Böling, D. (2006): Das CHE Forschungsranking deutscher Universitäten 2006. Arbeitspapier 79. Gütersloh. [http://www.che.de/downloads/CHE\\_Forschungsranking\\_2006.pdf](http://www.che.de/downloads/CHE_Forschungsranking_2006.pdf). [Last visited: 02/26/2010].
- Bundesministerium für Bildung und Forschung (BMBF) (2008): Bundesbericht Forschung und Innovation 2008. [http://www.bmbf.de/pub/bufi\\_2008.pdf](http://www.bmbf.de/pub/bufi_2008.pdf). [Last visited: 02/26/2010].
- Butler, L. (2006): Research assessment: moving beyond journal outputs. <http://www.sussex.ac.uk/Units/spru/events/ocs/viewpaper.php?id=96>. [Last visited: 08/14/2009].
- Butler, L. and Visser, M.S. (2006): Extending citation analysis to non-source items. *Scientometrics* 66 (2), 327-343.
- Deutsche Forschungsgemeinschaft (DFG) (2006): Förder-Ranking 2006. Institutionen – Regionen – Netzwerke. DFG-Bewilligungen und weitere Basisdaten öffentlich geförderter Forschung. Weinheim.
- Europäische Kommission – Eurostat (2008): Europa in Zahlen – Eurostat Jahrbuch 2008.
- European Science Foundation (ESF) (2008): Window to Science. Information Systems of European Research Organizations. Report of the EUROHORCs – ESF Working Group on a Joint Research Information System. Strasbourg.
- Hetmeier, H.-W. (1998): Methodik und Berechnung der Ausgaben und des Personals der Hochschulen für Forschung und Entwicklung ab dem Berichtsjahr 1995. *Wirtschaft und Statistik* 2/1998.
- Hornbostel, S. (2001): Third Party Funding of German Universities. An Indicator of Research Activity? *Scientometrics* 50 (3), 523-537.
- Hornbostel, S. (2006): Leistungsmessung in der Forschung. Von der Qualitätssicherung der Lehre zur Qualitätsentwicklung als Prinzip der Hochschulsteuerung. *Beiträge zur Hochschulpolitik* 1, 219-228.
- Hornbostel, S. (2007): Theorie und Praxis von Hochschulrankings. In: Statistisches Bundesamt (Ed.): Statistik und Wissenschaft, Band 11. Amtliche Hochschulstatistik und Hochschulrankings, 6-13.
- Hornbostel, S. (2008a): Gesucht: Aussagekräftige Indikatoren und belastbare Datenkollektionen. Desiderate geisteswissenschaftlicher Evaluierung in Deutschland. In: Lack, E. and Marksches, Ch. (Eds.): What the hell is quality? Frankfurt/New York.
- Hornbostel, S. (2008b): Bologna und die Forschung. In: Rudinger, G./Krahn, B. and Rietz, Ch. (Eds.): Evaluation und Qualitätssicherung von Forschung und Lehre im Bologna-Prozess. Bonn.
- Hornbostel, S. and Heise, S. (2006): Die Rolle von Drittmitteln in der Steuerung von Hochschulen. In: Berthold, Ch. (Ed.): Handbuch Wissenschaftsfinanzierung. Berlin.
- Jaeger, M. (2006): Leistungsorientierte Budgetierung: Analyse der Umsetzung an ausgewählten Universitäten und Fakultäten/Fachbereichen. In: HIS (Ed.): Kurzinformation, A 1/2006.
- Jansen, D. (Ed.) (2007): New Forms of Governance in Research Organizations – Disciplinary Approaches, Interfaces and Integration. Dordrecht.

- Michelson, E.S. (2006): Approaches to research and development performance assessment in the United States: an analysis of recent evaluation trend. *Science and Public Policy* 33 (8), 546-560.
- Moed, H.F./Glänzel, W. and Schmoch, U. (Eds.) (2005): *Handbook of Quantitative Science and Technology Research. The Use of Publication and Patent Statistics in Studies of S&T Systems*. Dordrecht.
- National Science Board (2006a): *Science and Engineering Indicators. Volume I*, <http://www.nsf.gov/statistics/seind06/pdf/volume1.pdf>. [Last visited: 02/26/2010].
- National Science Board (2006b): *Science and Engineering Indicators. Volume II*, <http://www.nsf.gov/statistics/seind06/pdf/volume2.pdf>. [Last visited: 02/26/2010].
- Organisation for economic co-operation and development (OECD) (2008a): *OECD Factbook 2008. Economic, Environmental and Social Statistics*.
- Organisation for economic co-operation and development (OECD) (2008b): *OECD in Figures 2008*.
- Schmoch, U./Rammer, Ch. and Legler, H. (2006): *National Systems of Innovation in Comparison*. Dordrecht.
- Statistisches Bundesamt (2008): *Datenreport 2008*. Wiesbaden. [http://www.destatis.de/jetspeed/portal/cms/Sites/destatis/Internet/DE/Navigation/Publikationen/Querschnittsveroeffentlichungen/Datenreport\\_\\_downloads,templateId=renderPrint,psml\\_\\_nnn=true](http://www.destatis.de/jetspeed/portal/cms/Sites/destatis/Internet/DE/Navigation/Publikationen/Querschnittsveroeffentlichungen/Datenreport__downloads,templateId=renderPrint,psml__nnn=true). [Last visited: 02/26/2010].
- Vahl, Ch.-F. (2008): *Forschungsförderung durch leistungsorientierte Mittelvergabe (LOM): Argumente für eine medizinische Wissenschaftskultur jenseits der Impact-Punkte*. *Zeitschrift für Herz- Thorax- Gefäßchirurgie* 22, 94-97.
- Wissenschaftsrat (WR) (2008): *Pilotstudie Forschungsrating. Empfehlungen und Dokumentation*. Cologne.



## 4. ECONOMY AND LABOR MARKETS

### 4.1 Data from the Federal Employment Agency

*Stefan Bender and Joachim Möller*

Contact:

Stefan Bender  
Institute for Employment Research (IAB)  
Regensburger Strasse 104  
90478 Nuremberg  
e-mail: stefan.bender[at]iab.de

## Abstract

The supply of German microdata for labor market research has been rapidly growing in recent years. This paper reports on this development, focusing on three key aspects: the establishment of Research Data Centers, the creation of new anonymization techniques for establishment data and the spate of scientific analyses and evaluation studies of active labor market programs based on administrative data in the aftermath of labor market reforms, especially the new Social Code II. Substantial progress has been made in all these fields with respect to the availability of adequate data, for instance, through the combination of different datasets. However, there is still large room for improvements. For future development, we recommend to focus on three primary areas: (1) the influence of researchers on data production; (2) (internationally) combinable datasets, and (3) the establishment of an international infrastructure for data access.

Keywords: labor market, data access, administrative data, linked employer-employee data, Research Data Center, evaluation of labor market programs, Social Code II

## 1. Introduction

There are numerous reasons for collecting labor market data. One could stress the importance of having reliable data for research on changes in the wage and employment structure, or one could point to the fact that labor market policies require evaluation studies for selecting the most effective of efficient instruments. Since this is a report on data, and our aim is a description of the German information infrastructure and its development we do not intend to go into more details here. Instead, we will restrict ourselves to referring to the illuminating introduction in Dan Hamermesh's article, "Fun with Matched Firm-Employee Data: Progress and Road Maps." In this article, Hamermesh raises the question,

"What generates scientific progress (assuming that we can use the term science to talk about economics)? Does it matter whether causation runs from data to problems or from problems to data? I think it does" (2008: 664).

We agree with Hamermesh that without datasets like the Panel Study of Income Dynamics (PSID) in the US or the German Socio-Economic Panel (SOEP, *Sozio-oekonomisches Panel*) our knowledge about the intertemporal labor supply or the intergenerational transmission of inequality, for instance, would be very limited.<sup>1</sup> Without the availability of administrative data for the

---

1 "In terms of issues of firm behavior in particular, I doubt that we would even have thought about the issues in the way we now do without the availability of this type of data. No doubt the opposite is also often true, but my purpose here is to talk about the former. To

evaluation of active labor market programs, we would know hardly anything about their effects. Kluve (2006), for example, reports that almost 80 percent of all microeconomic evaluation studies in Europe are based on administrative data. Without the relevant microdata on establishments, the work of, say, Dunne et al. (1989) or Davis and Haltiwanger (1992), would not have had the powerful influence that it has actually had.

Writing this motivational section during the first half of 2009, one cannot ignore the financial crises and its possible effects on the labor market. The World Recession triggered by the financial crisis hit economies surprisingly and had no predecessor in recent times. Moreover, it is an international phenomenon which requires a cross-border perspective. For such a situation the empirical basis is weak. Because of the lack of adequate data, deeper analyses are scarce and most researchers are more or less speechless when it comes to the concrete consequences of the crisis on firms' behavior and labor market outcomes.

Due to its rapid development in recent years, the German data infrastructure for the labor market has reached a satisfactory stage of completion. What we can learn from the current crisis, however, is that we need internationally comparable datasets at the micro-level and – more and more – a combination of these. These datasets should cover all important spheres of content, such as trade, foreign direct investments, offshoring, outsourcing, labor flows, earnings, strategic planning etc.

## 2. The situation before the 2001 KVI report<sup>2</sup>

Compared to the situation as it existed some years ago, access to confidential microdata – often critical for labor market research – has improved considerably. Prior to the 2001, the German Commission on Improving the Information Infrastructure between Science and Statistics, the creation of the FiDASt network “Firm-Level Data from Official Statistics” (*FirmenDaten aus der Amtlichen Statistik*), and the so-called *Schalterstelle* of the Institute for

---

paraphrase Matthew 5, ‘Blessed are the data developers because they inspire the creation of knowledge.’ Creating data is a very thankless task for which one gets very few points. Yet so many of our ideas are inspired by new data, and so much of research rests on innovations in questions and data collection that are barely, if at all, acknowledged by the more technical researchers” (Hamermesh 2008: 664).

2 Our report will focus primarily on the current situation, developments since KVI in 2001, and the short-term developments needed in the Federal Employment Agency (BA, *Bundesagentur für Arbeit*) data and the Institute for Employment Research (IAB, *Institut für Arbeitsmarkt- und Berufsforschung*) data. The other report in this volume focusing on labor market data, by Hilmar Schneider, will elaborate on the situation outside the BA and IAB.

Employment Research (IAB, *Institut für Arbeitsmarkt- und Berufsforschung*) that provides remote data access, represented milestones in labor market data availability. The FiDASt project was based on the need for access to data from the various Statistical Offices of the German *Länder*. A network was created for analyzing firm-level data. In some cases, remote data access could be used, in others the researcher was given the status of an unpaid employee at the Statistical Office (see Wagner in this publication for more details).

At the time of the 2001 KVI report, the only available microdata for labor market researchers were the IAB Employment Sample and the German Microcensus. In their advisory report to the Commission, Falk and Steiner (2000) argued that additional micro datasets should be created for the scientific community including information on active labor market policies, the IAB linked employer-employee dataset, or aggregated information on establishments from employment statistics. Information on marginal employment and forms of work outside the standard working-hour arrangements, being typical in Germany (such as temporary employment, temporary contracts, and self-employment), should be improved and made available, too. The high demand for establishment and firm-level data should be taken into account. Since quantitative information on establishments and firms is hard to anonymize, effort has to be put into investigating how these data can be offered to researchers. Basic information on employment and income should be made available via the Internet for free (especially employment and unemployment rates by skill level and differentiated by age and sex). When compared to the data provision of the US Bureau of Labor Statistics, the information provided by the Federal Statistical Office and the Federal Employment Agency (BA, *Bundesagentur für Arbeit*) seemed inadequate and left much room for improvement.

### **3. Situation since the 2001 KVI report**

#### *3.1 The establishment of Research Data Centers and Data Service Centers in Germany*

Following the suggestions of the 2001 KVI report, most of the important data producers in Germany – responsible for firm-level, organizational, and labor-market data as well as household income, poverty, and wealth data – have created Research Data Centers. Since 2000, one of the landmark developments in the improvement of the German data infrastructure was the establishment of four publicly funded Research Data Centers: the Research Data Center of the BA at the IAB, the Research Data Center of the German Pension Insurance (RV, *Deutsche Rentenversicherung*), and the Research Data Centers of both the Federal Statistical Office and the Statistical Offices of the

German *Länder*. Two Data Service Centers – the German Microdata Lab at the Center for Survey Research and Methodology (ZUMA, *Zentrum für Umfragen, Methoden und Analysen*) and the International Data Service Center at the Institute for the Study of Labor (IZA, *Forschungsinstitut zur Zukunft der Arbeit*) – were also established (see Bender et al. in this publication for more details).

The Research Data Center of the BA at the IAB was created in April 2004 and its micro datasets include the IAB Establishment Panel, the IAB Employment Sample (IABS), the BA Employment Panel (BAP), the Integrated Employment Biographies Sample (IEBS), the Establishment History Panel (BHP), the linked employer-employee dataset from the IAB (LIAB), the cross-sectional survey “Life Situation and Social Security 2005” (LSS 2005), and the first wave of the panel study “Labor Market and Social Security” (PASS, *Panel “Arbeitsmarkt und soziale Sicherung”*).<sup>3</sup>

### 3.2 New developments in anonymization techniques

In recent years, the public demand for microdata has increased dramatically. But statistical agencies face a dilemma. Although they might be willing to provide all the information required, it might not be possible to release these datasets for reasons of confidentiality. The natural desire to enable as much research as possible with collected data must take a back seat to the confidentiality that is guaranteed to the survey respondent. When confidentiality comes into play, potential respondents might be less willing to provide sensitive information, might intentionally provide wrong answers, or might even be unwilling to participate at all, with devastating consequences for the quality of the data collected (Lane 2005).

For this reason, there has been a variety of methods developed to provide as much information to the public as possible while satisfying the access restrictions needed to maintain the quality of the collected data (Willenborg and de Waal 2001; Abowd and Lane 2004). For German establishment datasets, a broad literature for anonymization approaches has developed, mostly based on perturbation techniques (e.g., Brand 2002; Gottschalk 2005; Rosemann 2006; Drechsler et al. 2008).

Official statistics in Germany, together with researchers using their data, carried out a research project on the “Factual Anonymization of Business Microdata,” which was finished in summer 2005. For the project, a large amount and variety of perturbation approaches were tested. Moreover, test analyses with anonymized real data were undertaken and the results were

---

3 Documentation in English can be found on the webpage for nearly all datasets. Descriptions are also published in the “Data Watch” section of *Schmollers Jahrbuch*.

compared with those obtained from analyses based on original data.<sup>4</sup> Several anonymized cross-sectional data were made available. They included data on cost structures in industry and in retail trade, as well as profit-tax data. Detailed descriptions of the Scientific Use Files can be found in Lenz et al. (2005), Vorgrimler et al. (2005), and Scheffler (2005). Recently, similar approaches have been used to anonymize business statistics. These include the German data of the Continuing Vocational Training Survey 1999 and the German Structure of Earnings Survey 2001.

The aim of the follow-up project “Business Statistical Panel Data and Factual Anonymization” (FAWE-Panel, *Wirtschaftsstatistische Paneldaten und faktische Anonymisierung*) was to improve researcher access to panel data from official statistical offices and the BA. In cooperation with different Research Data Centers and the Institute for Applied Economic Research (IAW, *Institut Arbeit und Wirtschaft*), the objective was:

- to expand the data supply for researchers by adding individual business statistical panels,
- to optimize the potential of business panel data for analysis, and
- to research the possibility of the factual anonymization of panel data in the field of economic statistics with the goal of ultimately making these available as Scientific Use Files.

Results from this project are published in a special volume of *Advances in Statistical Analysis* (Pohlmeier and Ronning 2008), and a special volume of *Wirtschafts- und Sozialstatistisches Archiv* (Bender et al. 2009).

While anonymization techniques have helped to make establishment and firm-level data more accessible, the degree of anonymization for some individual data has decreased over time. For example, panel anonymization was abandoned for the IAB Employment Sample.

### 3.3 A new and unexpected situation: *The Social Code II*<sup>5</sup>

With the deep changes to the German system of unemployment insurance and welfare benefits that took place through the labor market reforms in 2005 (Social Code II / SGB II), a new data infrastructure was implemented in Germany. These changes occurred in three main areas:

- (1) Job search and participation in active labor market schemes
- (2) Data stemming from the SGB II software, A2LL
- (3) Data from 69 districts where local authorities (zkT, *zugelassene kommunale Träger*) are responsible for administering the unemployment assis-

<sup>4</sup> For more information about this project see Lenz et al. (2006).

<sup>5</sup> This section is taken from Koch et al. (2008). It is shortened and translated.

tance (ALG II, *Arbeitslosengeld II*). Data are available through the BA via the interchange program XSozial.

With the start of the SGB II, the BA expanded its IT administrative procedures. In July 2006, for example, the BA introduced an integrated program for occupational counseling and an employment service. The change in the procedures temporarily affected the quality of the data in the years 2006–2007.

Because the reform proceeded so rapidly, A2LL was very hastily implemented in 2004. As a result, a number of problems occurred. In addition to difficulties with the software, which made headlines in all the German newspapers, there was no interface between the data produced by A2LL and all the other datasets belonging to the BA and stored in the BA data warehouse. To integrate these data is time-consuming and costly. In 2005 and 2006, the A2LL data changed significantly due to the synchronization of individuals over time and due to the introduction of the concept of so-called “communities of need” (*Bedarfsgemeinschaften*) in the SGB II. This concept implies that individuals in disadvantaged households take responsibility for each other and, therefore, entails a different logic of representativeness of the data. Before SGB II, individuals were comprised in the data when they received unemployment insurance. Since SGB II, unemployed individuals living together with a partner whose earnings are above a certain level are no longer eligible for receiving unemployment benefits. The situation changed also for low income earners. Before SGB II, they were not eligible to receive additional money from unemployment insurance; after SGB II, this became possible.

The situation of the data from XSozial is rather complex. Because of the initial problems encountered by all participating institutions, there are no microdata available for the first few years. At the end of 2006, the BA began working to build up an adequate micro database. To date, this work is not finished and thus the data is not available for researchers. The main reasons for the delay are: missing data, the different time structure, and the different definitions and collection of variables in the different software systems. Because of the lack of information in the XSozial data, a harmonized dataset will end up with only a few basic variables.

At the IAB, SGB II data will be integrated into the Integrated Employment Biographies. At the moment, this is the dataset where employment information based on the social security system, receipt of benefits, participants in active labor market policy measures, and job seekers are integrated over time (since 1990) and harmonized.

Information on unemployment benefits between 2005 and the fall of 2007 can be distinguished between traditional unemployment benefits and the “new” Unemployment Benefit II (ALG II). Not integrated into XSozial are data from the 69 districts where local authorities are responsible for administering Unemployment Benefit II.



Everyone should be aware that – beginning in 2005 – Germany has some white regions on data maps where previously there had been information. There are no microdata covering the years 2005 and 2006 available for the 69 districts. This deficiency extended to statistics covering social benefits. As an unfortunate result there are no administrative microdata available covering one of the biggest changes to the labor market in German history (2004–2005).

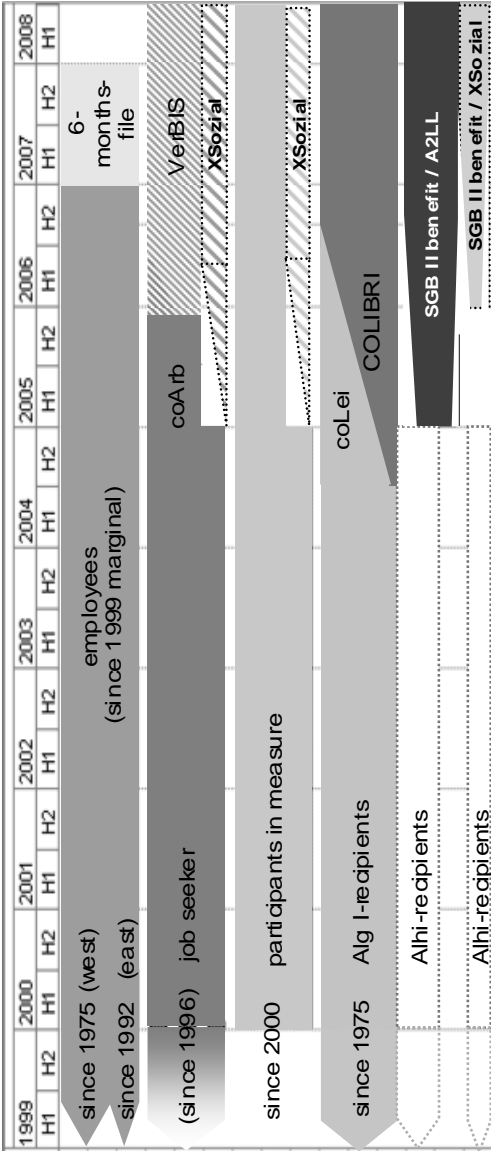
The SGB II has introduced different possibilities for organizing social and unemployment assistance at the regional level – such as *Arbeitsgemeinschaften* (ARGE) – a joint organization between local authorities and the BA – or *Kreise mit geteilter Trägerschaft* where the different forms of assistance are separated. Due to these different organizational forms, non-compatible software for job seekers and individuals in need under the SGB II has been used. Hence, the situation does not seem to be improving. Since 2006, nearly everyone has been using the same software (A2LL). All rescue efforts for gathering microdata for those white regions for 2005 have failed. XSozial data for 2005 will not be available if no techniques like imputation are used to replace missing data.

In addition to administrative data and “standard” surveys, two new survey tools can be used for analyzing the situation of households in need or individuals living in those households. The cross-sectional survey, “Life Situation and Social Security 2005” (LSS 2005) covers the period 2005–2006 and asked 20,832 recipients about unemployment benefits. It is a unique survey because it covers the time during which the Social Code II was first introduced. The panel study “Labor Market and Social Security” started its first wave in 2006–2007 and surveys the households of recipients receiving support from Unemployment Benefit II as well as low wage earners. The panel covers 12,794 households and 18,954 individuals. Both datasets are available for researchers via the Research Data Center of the BA at the IAB.

### *3.4 The new spirit: labor market research with administrative data*

Until the late 1990s, the evaluation of active labor market policies was in a state of hibernation. Starting with the so-called Hartz Reforms at the end of 2002, the German Bundestag commissioned an evaluation of these reforms as a large research project (around 100 researchers in nearly twenty institutions with a budget of around 10.3 million Euros directed into different projects). These investigations into the evaluation of active labor-market policy were extensive, and point to a new quality in the discourse around policy and research in Germany. Based on empirical evidence, labor-market policy is now conceived as something like a “learning system.” That is, new instruments are temporarily introduced and subsequently evaluated by researchers, which leads to im-

Figure 1 : Short summary of the gaps in administrative data stemming from SGB II



proved instruments (if there is a need for improvements). In this framework, labor-market policy acts as a pioneer for other policy areas. This prescribes an important role for research (Social Code II and III) and to have a detailed, well-defined mandate. The evaluation of the so-called Hartz Reforms thus represents a significant and singular project that will influence future labor-market research in Germany and that exemplifies a new form of research-based policy advice (Heyer 2006).

Another milestone was the research network, “Flexibility of Heterogeneous Labor Markets” financed by the German Research Foundation (DFG, *Deutsche Forschungsgemeinschaft*). The network was established in 2004 with the objective of finding solutions to the permanent challenge of responding quickly and effectively to changes in the labor market. A central aim of the research program is thus to analyze rigidities that have developed historically in labor-market institutions and to investigate whether and to what extent deregulation is possible and/or necessary. The program has organized a research group to tackle the pertinent questions by actively drawing on data from the Research Data Center of the BA at the IAB as well as from other Research Data Centers. This research network and the Research Data Center of the BA at the IAB are also active in a more informal network and have jointly stimulated data production processes and higher data quality. The fruit of this work are several published articles in international journals.<sup>6</sup>

In addition to these networks, there are many researchers interested in labor market use of data from the Research Data Center (see Bender et al. in this publication for more details). As an illustration, in September 2008 there were already 341 publications drawing from data made available through the Research Data Center of the BA at the IAB. Twenty-seven of these were publications in SSCI journals and an additional forty-two appeared in peer-reviewed journals. Judging from this and the number of publications in process as indicated by an increasing number of discussion papers – Germany is no longer a wasteland for empirical labor-market research. To the contrary, a quantitative and a qualitative improvement is evident in papers presented at international conferences based on BA-IAB data, as well as those published in international, high-quality journals. There should and could be more done to improve research with German datasets, but it should be kept in mind that the starting point for significant change only dates to 2004, and that research with datasets is a slow process of diffusion. There must be some visible articles before other researchers begin working with the data.

---

6 See [www.zew.de/dfgflex](http://www.zew.de/dfgflex) for further information.

#### 4. What is missing?

In the wake of these developments over the past few years, some of the gaps mentioned by Falk and Steiner in 2000 have been closed while others have not. For example, the availability of establishment data, data for evaluating active labor-market policies, linked employer-employee data, and data for specific groups in the labor market are now available. But precise and comprehensive data for earnings and wages are still not available. Neither the employment statistics nor the German Microcensus have changed their concepts over the last few years, for example. An exception to this is the SOEP, which has raised the quality of its income and assets data (Anger et al. 2008).<sup>7</sup> Although some specific groups in the labor market, like the marginally employed, have been incorporated into the administrative microdata of the BA-IAB, there is still no information available for temporary employment, temporary contracts, or self-employment in Germany.<sup>8</sup>

Basic aggregated information for some labor market indicators (such as employment) is available via the Internet for free, but the information system in Germany still lags behind official statistics in other countries, such as the US Bureau of Labor Statistics. The information given by the Federal Statistical Office and the BA could be improved. Basic aggregated information such as wages or regional price indices are still missing.<sup>9</sup>

#### 5. Outlook

Because researchers use administrative data more frequently, there is now a demand to change the production of administrative data to suit researcher needs in two respects: (a) to add additional variables like working hours, contract type, or anonymized case worker IDs to existing datasets; and (b) to provide more information about data-generating processes, because the quality of administrative data is an underdeveloped research field. Some planned changes in the development of variables are “automatically” benefitting researchers – like the addition of working hours or the inclusion of an internationally comparable occupation code in employment statistics (Stegmann 2009). However, the research community needs to exert constant pressure to make more relevant microdata from administrative processes available as well as to enhance data quality.

---

7 See the paper of Schneider in the same book for a deeper discussion of the topic.

8 The backbone of the administrative data in the BA / IAB will be changed in this direction (Stegmann 2009).

9 The RatSWD initiated expertises on regional price indices in 2008.

It is increasingly important to create reliable and precise microdata for topics of current relevance (e.g., innovation, globalization) and to ensure that it is available for research. For example, a “double” linked employer-employee dataset could be used for the project, “Further Training as a Part of Lifelong Learning,” co-funded by the Leibniz Association. For the first time in Germany, data stemming from surveys and administrative data for both employer and employee groups are combined and available (Bender et al. 2008). Combining the available data (administrative data, survey data, commercial data, and Internet data) is extremely important. However, because we do not have comparable unique identifiers in German datasets, we need more research on record linkage techniques and also to engage in dialogue with representatives of data protection and with legislators (see the papers by Schnell and Metschke in this publication).

The projects Official Firm Data for Germany (AFiD, *Amtliche Firmendaten für Deutschland*) and Combined Firm Data for Germany (KombiFiD, *Kombinierte Firmendaten für Deutschland*) will extend the range of data in two directions: AFiD will integrate economic and environmental data from official statistics, and KombiFiD will link company data from official statistics, the Deutsche Bundesbank and the BA-IAB for the first time (Hethy and Spengler 2009). The project Biographical Data of Selected Social Insurance Agencies in Germany (BASiD, *Biografiedaten ausgewählter Sozialversicherungsträger in Deutschland*) will combine administrative data from the German Pension Insurance with IAB data. It will offer a Scientific Use File for researchers in Germany and abroad as well as on-site use. International datasets are needed because individuals are migrating and many firms no longer remain within national borders. Datasets should not have those restrictions either.<sup>10</sup>

It is becoming increasingly important to establish an international infrastructure for data access (including translation, harmonization, integrated metadata systems, integrated access, and remote access) and to coordinate the different developments taking place in different organizations and/or countries.

---

10 “However, an improved statistical infrastructure is needed not only on the national level. As the European research landscape evolves, it produces increased demands on the data infrastructure in order for the social sciences and economics to develop their full potential in the area of social comparisons as well. By actively participating in important developments at both the national and international levels, the RatSWD intends to work even more intensively in this important field in the future. It already provides a platform for a fundamental discussion and planning process that is almost one of a kind both in Europe and beyond. If international and interdisciplinary strategic planning is to be successful in fostering empirical research and improving the research infrastructure, however, greater involvement of the professional scientific organizations representing the social sciences and economics will be urgently needed. The RatSWD will endeavor to promote this involvement” (Solga and Wagner 2007: 4).

Through the activities of the German Data Forum (RatSWD) itself, German researchers and the Research Data Centers have started to present the German model of data access and data infrastructure to the international research community. The German experience of organizing access and building up an infrastructure could offer a blueprint for how an equivalent international system might be established. There is a need for both coordination and advocacy in this area to press for the activities necessary for its realization and to steer them in the right direction.

## References:

- Abowd, J.M. and Lane, J. (2004): New Approaches to Confidentiality Protection: Synthetic Data, Remote Access and Research Data Centers. *Privacy in Statistical Databases*. New York, 282-289.
- Anger, S./Bowen, D.A./Engelmann, M./Frick, J.R./Goebel, J./Grabka, M.M./Groh-Samberg, O./Haas, H./Headey, B./Holst, E./Krause, P./Kroh, M./Kurka, C./Lohmann, H./Pischner, R./Rahmann, U./Schmitt, C./Schräpler, J.-P./Schupp, J./Sieber, I./Siedler, T./Spieß, C.K./Spieß, M./Tucci, I. and Wagner, G.G. (2008): 25 Wellen Sozio-oekonomisches Panel. *Vierteljahrshefte zur Wirtschaftsforschung* 77 (3), 9-14.
- Bender, S./Rosemann, M./Zühlke, S./Zwick, M. (2008): Betriebs- und Unternehmensdaten im Längsschnitt – neue Datenangebote und ihre Forschungs-potenziale. Vorwort. *Wirtschafts- und Sozialstatistisches Archiv*. AStA 2, 189-191.
- Bender, S./Fertig, M./Görlitz, K./Huber, M. and Schmucker, A. (2009): WeLL – unique linked employer-employee data on further training in Germany. *Schmollers Jahrbuch. Zeitschrift für Wirtschafts- und Sozialwissenschaften* 129, 637-643.
- Brand, R. (2000): Anonymität von Betriebsdaten – Verfahren zur Erfassung und Maßnahmen zur Verringerung des Reidentifikationsrisikos. *Beiträge zur Arbeitsmarkt- und Berufsforschung* 237.
- Brand, R. (2002): Masking through Noise Addition. *Inference Control in Statistical Databases*. Berlin/Heidelberg, 97-116.
- Brand, R./Bender, S. and Kohaut, S. (1999): Possibilities for the creation of a scientific-use file for the IAB-Establishment-Panel. *Statistical Data Confidentiality Proceedings of the Joint Eurostat/UN-ECE Work Session on Statistical Data Confidentiality Held in Thessaloniki in March 1999*. Eurostat. Brüssel, 57-74.
- Davis, S. and Haltiwanger, D. (1992): Gross Job Creation, Gross Job Destruction and Employment Reallocation. *Quarterly Journal of Economics* 107, 819-863.
- Drechsler, J./Dundler, A./Bender, S./Rässler, S. and Zwick, T. (2008): A new approach for disclosure control in the IAB Establishment Panel – multiple imputation for a better data access. *Advances in Statistical Analysis* 92, 439-458.
- Dunne, T./Roberts, M. and Samuelson, L. (1989): The Growth and Failure of U.S. Manufacturing Plants. *Quarterly Journal of Economics* 104, 671-698.
- Falk, M. and Steiner, V. (2000): Informationsfeld Erwerbstätigkeit/Einkommen im KVI-Gutachten.
- Gottschalk, S. (2005): Unternehmensdaten zwischen Datenschutz und Analysepotenzial. *ZEW Wirtschaftsanalysen* 76. Baden Baden.
- Hamermesh, D.S. (2008): Fun with Matched Firm-Employee Data: Progress and Road Maps. *Labour Economics* 15 (4), 663-673.
- Hethey, T. and Spengler, A. (2009): Matching process generated business data and survey data – the case of KombiFiD in Germany. *FDZ Methodenreport* 01/2009 (en).
- Heyer, G. (2006): Zielsetzung und Struktur der "Hartz-Evaluation". *Zeitschrift für ArbeitsmarktForschung* 39 (3/4), 467-476.

- Kluve, J. (2006) The Effectiveness of European Active Labor Market Policy. IZA Discussion Paper 2018.
- Koch, S./Kupka, P. and Steinke, J. (2008): Aktivierung, Erwerbstätigkeit und Teilhabe – Die Wirkungen der Grundsicherung für Arbeitsuchende (Ergebnisse der Wirkungsforschung nach §55 SGB II von 2005 bis 2008).
- Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik (KVI) (Ed.) (2001): Wege zu einer besseren informationellen Infrastruktur. Baden-Baden.
- Lane, J. (2005): Optimizing the Use of Microdata: An Overview of the Issues. Paper presented at the Joint Statistical Meetings. <http://client.norc.oregonstate.edu/jole/SOLEweb/Accessmicrodata%5B1%5D.pdf>. [Last visited: 02/10/2010].
- Lenz, R./Vorgrimmler, D. and Rosemann, M. (2005): Ein Scientific-Use-File der Kostenstrukturerhebung im Verarbeitenden Gewerbe. *Wirtschaft und Statistik* 2, 91-96.
- Lenz, R./Rosemann, M./Sturm, R. and Vorgrimmler, D. (2006): Anonymising Business Micro Data – Results of a German Project. *Schmollers Jahrbuch* 126 (4), 635-651.
- Pohlmeier, W and Ronning, G. (2008): Microeconometrics and disclosure control. *Advances in Statistical Analysis* 92, 351-357.
- Rosemann, M. (2006): Auswirkungen datenverändernder Anonymisierungsverfahren auf die Analyse von Mikrodaten.
- Stegmann, M. (2009): Änderungen der Erfassung der Angaben über Bildung, Beruf und Beschäftigungsform im Meldeverfahren zur Sozialversicherung. *Deutsche Rentenversicherung*, 487-500.
- Scheffler, M. (2005): Ein Scientific-Use-File der Einzelhandelsstatistik 1999. *Wirtschaft und Statistik* 3, 197-200.
- Solga, H. and Wagner, G.G. (2007): A Modern Statistical Infrastructure for Excellent Research and Policy Advice – Report on the German Council for Social and Economic Data during its First Period in Office (2004 - 2006). RatSWD Working Paper No. 2.
- Vorgrimmler, D./Dittrich, S./Lenz, R. and Rosemann, M. (2005): Ein Scientific-Use-File der Umsatzsteuerstatistik 2000. *Wirtschaft und Statistik* 3, 197-200.
- Willenborg, L. and de Waal, T. (2001): *Elements of Statistical Disclosure Control*. New York.



## 4.2 More and Better Data for Labor Market Research

### **Proposals for Efficient Access to the Currently Unused Potential of Official Statistical Data**

*Hilmar Schneider*

Contact:

Hilmar Schneider  
IZA – Institut zur Zukunft der Arbeit  
Schaumburg-Lippe-Str. 7-9  
53113 Bonn  
Germany  
e-mail: [schneider\[at\]iza.org](mailto:schneider[at]iza.org)

## 1. Introduction

The economic assessment of issues related to the labor market requires a sound and comprehensive database. A key variable for understanding labor market processes is hourly wages. For example, hourly wages can explain whether people work or not, and if they work, how much they work. The related earnings are a major source of income and hence a key to explaining consumption. Hourly wages also explain why and how long people remain unemployed as well as how much they invest in human capital. On the employer's side, hourly wages explain the number of workers being hired or dismissed, the location of firms, and the optimal form of specialization. Last but not least, hourly wages have a crucial impact on income distribution. Given the importance of hourly wages, it appears more than surprising that official statistics provide little data support for this central part of the economy. With comparably little effort, significant improvements could be made to the utilization of existing data. However, it would require the support of the legislative. On the one hand, this affects the conception of official surveys, and on the other hand, it requires a legal basis for merging official microdata for scientific purposes.

The logic of the existing surveys is still rooted in an antiquated understanding of statistics that sees its objective in compiling aggregate figures with the aid of independent samples designed specifically for a particular purpose. Structurally related data like working hours and wages are collected by separate surveys. Hourly wages can therefore only be computed on an aggregate level. The analytical potential of such datasets will remain limited as long as it is impossible to create broad linkages with structural characteristics on the level of observation units.

If the samples were conceptualized to enable information on the firm level to be linked with information on the individual and household level, the analytical potential of the data would be significantly increased, while at the same time reducing the effort required for data collection. Indeed, it was for this reason that the 2001 KVI report called for legal regulations to provide a possibility for exact data linkage – without the express agreement of all respondents – for purely statistical purposes (see KVI 2001, recommendation 34).

Aggregated statistics are insufficient because they do not provide a basis for conclusive answers about the structural causes of differences or changes in aggregate indicators. In the scholarly literature, this phenomenon is known as an ecological fallacy (Robinson 1950). Typical questions of labor market research can only be answered based on microdata, in which micro-level heterogeneity in observation units – which, for labor market research, consist mainly of individuals, households, and firms – serves as the identifying moment. Here, two kinds of heterogeneity are relevant: heterogeneity

between observation units on the micro-level (between variance) and heterogeneity over time (within variance). If the process under examination is a stationary one, and if characteristics with explanatory relevance are observable, it is sufficient to collect cross-sectional data; that is, to examine the heterogeneity between units of observation. Going beyond this to examine panel data is useful, first, when cohort effects are present, and second, when the observability of the characteristics with explanatory relevance is limited, which can essentially never be ruled out. With the aid of panel data, the potentially distorting effect of unobservable heterogeneity can be neutralized (see Baltagi 1995).

From this, it follows that there is a basic need to collect panel data on the micro-level. Numerous such data are already collected by statistical offices or originate from administrative processes. But there is still, to some extent, a lack of suitable means for researchers to access these data, and in some cases it would require that existing survey concepts be linked in a coherent manner.

The related request is not completely new. In an advisory report on employment and earnings for the 2001 KVI report, Falk and Steiner (2000) discussed the problems outlined above and formulated a series of recommendations to address them. The present chapter takes up these recommendations and assesses the hitherto achieved progress. Building on this, areas will be identified in which further action is needed. The focus here will be on the proposal for coordinated sampling procedures that would allow linkage of individual and firm-related data and can thus be seen as an extension of the recommendations by Falk and Steiner.

## **2. Prime example: Hourly wages**

As already mentioned, hourly wages are the key driver for empirically based economic analyses of the labor market. This requires information about paid salaries as well as information about the related working time. However, in official statistics, information on these variables is not only collected in separate surveys, but also in an insufficient manner.

Respondents to the Microcensus, for example, are surveyed regarding hours worked, but not with respect to the wages earned. They are asked instead just to provide personal net income or household net income. Since these figures can also include unknown amounts of transfers and other forms of income, it is impossible to calculate hourly wages. Approximately one-sixth of all German households receive transfers, such as housing allowances and unemployment benefits (see Rudolph 2008). According to the Federal Statistical Office, almost one quarter of households receive child allowances or parental allowances. Additionally, the tax and transfer system provides

very low wage earners with a major source of income compensation (BMAS 2008), meaning that the use of net income to approximate earned income will probably lead to completely erroneous conclusions.

Since 2003, the Income and Consumption Survey (EVS, *Einkommens- und Verbrauchsstichprobe*) has included survey questions not just on detailed income categories, but also on weekly working hours – albeit only the contractually stipulated working hours. Since overtime hours are also included in income earned, actual hourly wages are overestimated when dividing earned income by contractually stipulated working hours. Furthermore, the EVS is only conducted once every five years and is limited to specific household types (on this and other limitations, see Hauser 2010). In addition, the EVS is conducted on a voluntary participation basis, which can be accompanied by systematic selection bias, adversely affecting the representativeness of the survey further.

The Time Budget Survey carried out in 1991/92 and 2001/02 does not overcome the limitations mentioned above. Although the Time Budget Survey does contain a detailed record of time use over the course of the day, it is impossible to link this with earned income.

### **3. The recommendations of Falk and Steiner**

Official statistical agencies possess a series of micro datasets that are relevant to labor market research. In this context, Falk and Steiner (2000) examine the following datasets:

- (1) European Household Panel (EHP)
- (2) Microcensus
- (3) Income and Consumption Survey
- (4) Time Budget Survey
- (5) Salary and Wage Structure Survey
- (6) Official Industry Statistics on Firms and Companies
- (7) Social Security Statistics
- (8) Cost Structure Survey

The report concludes with approximately 15 recommendations, eight of which are relevant to the datasets listed above. These recommendations are summarized in the following:

- (1) Develop the Microcensus into a dataset relevant to labor market issues by including income characteristics.
- (2) Develop the Income and Consumption Survey into an instrument for annual accounts and expand it to include data on working hours.

- (3) Collect Microcensus data over the course of the year.
- (4) Develop the Microcensus further as an access panel.
- (5) Open up access to data for interested researchers.
- (6) Provide individual-level microdata from official statistics in factually anonymized form.
- (7) Provide firm- and company-level data from official statistics in simple anonymized form.
- (8) Comparatively evaluate data access through computer centers and remote data processing facilities that are subject to federal data protection regulations in the framework of a pilot project.

The first two of these recommendations should be seen as crucial, and relate to the problem mentioned at the outset: that the existing microdata have been used relatively little in labor market research because central variables are not available. In the area of individual-level data, this is especially the case for hourly wages, and in the area of firm-level data, it is true for the decomposition of value added into its labor and capital components.

#### **4. An assessment of progress achieved since 2000**

To start with, on a positive note, the number of micro datasets of potential interest for labor market research has increased significantly with the establishment of the Research Data Centers of the Federal Statistical Office and the Statistical Offices of the German *Länder*. Nearly all of the data collected by official agencies are now available to interested researchers, either as Scientific Use Files (SUFs) or as onsite files. The demand to provide access to the data has thus been met to a large extent. But what still has not been adequately addressed, at least by official statistical agencies, is the need to provide controlled remote data processing – despite the emphasis on this point in the 2001 KVI report and the fact that this is an integral component of the newly founded Research Data Centers' work from the Commission's point of view (see KVI 2001, recommendation 29).

Along with Research Data Centers, two Data Service Centers have also been established on the KVI's recommendations. The task of these Data Service Centers is to provide special services to data users and data producers (see Schneider and Wolf 2008). The Data Service Center at GESIS (GMD, *German Microdata Lab* at GESIS, *Leibniz Institute for the Social Sciences*) has taken on the task of developing special aids in the documentation and use of official microdata. The International Data Service Center at the Institute for the Study of Labor (IZA, *Forschungsinstitut zur Zukunft der Arbeit*) aims to meet the needs of an international research

community. The International Data Service Center provides, on the one hand, metadata on micro datasets relevant to the labor market, and on the other, assistance to foreign researchers in accessing German microdata. One of its outcomes has been the development of a tool for controlled remote data processing that goes by the name JoSuA (Job Submission Application) and can be used as a prototype for use in Research Data Centers. Up to now, however, it is used only by researchers from abroad who want to use SUFs from official statistics for projects conducted in cooperation with IZA. The researchers send their questions for analysis by JoSuA to the IDSC, where the actual data access occurs, after the International Data Service Center has checked the input for compliance with data protection regulations. Then, after the output is checked by the International Data Service Center for compliance with data protection regulations – and if permissible – the data are sent back to the researcher who made the request.

Although the number of micro datasets made available by statistical offices has increased substantially, quantity is not equivalent to quality. The value added of, for example, company data on continuing education collected in the framework of the European Continuing Vocational Training Survey (CVTS) remains very limited since this survey does not collect data on individual-specific characteristics, nor does it allow for linkage of existing individual datasets. Much the same is true of the statistics on students or marriages, to cite only two examples.

In regard to the two primary demands of Falk and Steiner, the progress achieved so far can be described as modest. With the new Microcensus Law of 2005, an opportunity to expand the survey to encompass measures of earned income has been lost for the foreseeable future. Nevertheless, the EVS has been surveying contractually stipulated weekly working hours since 2003. Before then, the survey was limited to surveying the categories of part-time and full-time employment as well as marginal employment. As mentioned above, the issue of the EVS's adequacy for computation of hourly wages has by no means been solved satisfactorily.

With the new version of the Microcensus Law, the call to conduct the Microcensus over the course of the year has been met. Previously, the Microcensus data always referred to the situation on a single reference date in April of the survey year, so the Microcensus was unable to reflect major seasonal fluctuations in employment behavior. Following extensive pilot surveys in the years 2003 and 2004, as of January 2005, the Microcensus is now conducted at monthly intervals. As a byproduct, the pilot surveys have provided labor market research with an interesting new micro database on the International Labour Organization (ILO) employment status.

The proposal to further develop the Microcensus as a so-called "access panel" has also been taken up by official statistical agencies. Since 2004, participants in the Microcensus have been asked for their agreement to

participate on a voluntary basis in further official statistical surveys (see Körner et al. 2006). The advantages of this procedure consist of the possibility to draw random samples with great flexibility and the ability to achieve higher response rates than is possible with the usual sampling procedures. In addition, the adoption of this procedure has led to not having to collect the same already available data repeatedly. This could potentially overcome the limitations identified above in the separate surveying of related characteristics. In practice, however, the potential here has by no means yet been exhausted, since only a relatively small portion of the variables surveyed in the Microcensus are actually included in the corresponding master file. For example, among the variables on working hours, only normal weekly working hours are included in the master file. Linkage with the EVS, for instance, offers no new information for the generation of hourly wages.

The selection of potential survey participants based on the Microcensus is mainly utilized in sample selection for the European Statistics on Income and Living Conditions (EU-SILC). This panel survey, conducted for the first time in 2005, is the successor to the European Community Household Panel (ECHP), which was carried out until 2001. Unfortunately, this survey also passed up its chance to create a solid database on individual hourly wages. Although detailed information is collected on current working hours, in the area of income, only data on monthly net income is collected, without differentiation into income components. Differentiated income data is only collected retrospectively for the previous year, but without corresponding data on working hours. Although it is conceivable to link current data on working hours with retrospectively collected earnings data at a later point in time, this is only useful for individuals with a continuous employment history. For those with career interruptions, or a change of their employer, retrospective annual income data do not match current working time data.

Even if little progress has been made so far in generating individual hourly wages based on official microdata, a development is currently underway that gives reason for hope. Here, efforts have been made to put the idea of merging microdata from various sources into practice as far as possible within the given legal constraints. This process has even succeeded in overcoming institutional boundaries, which should undoubtedly be counted as an accomplishment of the data infrastructure created so far. The effort that should be mentioned here above all is the project Combined Firm Data for Germany (KombiFiD, *Kombinierte Firmendaten für Deutschland*) (see also Möller and Bender 2010; Wagner 2010).<sup>1</sup> By linking firm-level data from the Federal Statistical Office, the Deutsche Bundesbank, and the Federal Employment Agency (BA, *Bundesagentur für Arbeit*), it aims at creating a combined database.

---

<sup>1</sup> See also: <http://www.kombifid.de>.



There is unquestionably vast potential in the coordinated linkage of previously independent surveys for improving and enhancing official statistics. An important step in this direction can also be seen in the conception of the 2011 Census, which plans to link the available information on the individual or household level from diverse registers (see Statistische Ämter des Bundes und der Länder 2004; Heinzel 2006). The procedure developed for this purpose should provide a valuable impetus toward a reconceptualization of coordinated sampling procedures in official statistics.

## 5. Resulting or remaining needs for action

A comparison of the data requirements described above and the progress achieved so far results in four basic recommendations for action that can be summarized briefly as follows:

- Provide controlled remote data processing in the Research Data Centers of the Federal Statistical Office and the Statistical Offices of the German *Länder*.
- Expand the program of household statistical surveys to include questions on hourly wages.
- Legally regulate exact data linkage – without the express agreement of all respondents – for purely statistical purposes.
- Conduct coordinated statistical surveys to create possibilities for data linkage.

The call for controlled remote data processing needs no further explanation. A tension exists between the demand to expand the programs of household statistical surveys to include questions on hourly wages and the recommendations that follow here. If it were possible to link different data sources with each other, the demand to expand the program of questions in these surveys would be superfluous. This demand should therefore be seen as a second-best solution, for it would only solve a specific – although very fundamental – problem. Coordinating statistical surveys to facilitate data linkage, on the other hand, would not only overcome the lack of data on hourly wages but would also solve many other problems. Without a sound legal foundation, however, this is not to be expected.

As long as such a fundamental legal solution remains a distant prospect, there is no alternative to expanding the survey program of the Microcensus and the EVS. The next opportunity to do so will arise with the harmonization of national household statistics that is planned on the European level. The

Microcensus Law will remain in effect only up to the year 2012. One year later, the law governing the EVS will expire. It is probable that legislators will decide to pass a comprehensive law reconceptualizing statistics oriented toward European guidelines. Already in November 2007, a workshop took place in Mannheim organized by GESIS together with the German Data Forum (RatSWD), the goal of which was to foster dialog between scholars and official statistical agencies on the programs of household statistical surveys. The Federal Statistical Office is addressing this question within the framework of the project "Reforming Household Statistics". Beyond this, a steering committee was appointed by the federal and state governments to advise on how to proceed. In order for a coordinated concept to be introduced into the legislative procedure in a timely manner, it would have to be agreed upon by the end of 2010 at the latest. On a European level, these considerations have found expression in the planning for a new ECHP. A pilot study was already carried out in 2008.

The alternative of coordinated linkage of data sources, on the other hand, would allow the information basis to be expanded much more effectively. It would permit improvement in the quality of the data collected, on the one hand, and would reduce the effort required to collect the data, on the other. The possibilities for linking data also would not have to be limited to surveys by the statistical offices. One could conceive, for example, of linking Microcensus samples with administrative data from the BA. While the BA data lack information on working hours, the Microcensus lacks information on earnings. When each is taken on its own, the datasets offer a limited basis for drawing conclusions. But combining them would result in a powerful basis for analysis in which, for example, the Microcensus data would compensate for the lack of information on working hours in the BA data.

This kind of strategy also goes far beyond the recommendations of Falk and Steiner (2000). While their recommendation was limited to complementary collection of individual or household data by official statistical agencies, the project KombiFiD demonstrates the possibilities that result from the coordinated linkage of firm-level data with individual- or household-level data beyond institutional boundaries. With KombiFiD, the BA plans to contribute only aggregated individual data, such as number of employees, age structure of employees, etc. In principle, however, this opens up the possibility for the reverse perspective: that of linking individual data with firm-level data and thus expanding in the direction of a linked employer-employee dataset.

In this case, the coordinated linkage of data sources requires that the sample of households and firms be drawn in a coordinated manner. This may go in two directions, which are not equally useful. It would be possible, for instance, to generate the Microcensus sample no longer just as a population sample, but also as a sample containing all employed persons from a firm

sample. The Microcensus sample, however, could not be drawn exclusively on this basis since otherwise, unemployed and non-employed persons would be excluded from the survey. It could therefore prove more sensible to take the reverse approach, using the employers of the Microcensus respondents as the basis for a firm sample.

Even if the perspective outlined here is of a more long-term nature, there is no reason to dismiss it as unrealistic. The reconceptualization of the 2011 Census as a register-based census has shown a promising way to overcome the basic obstacles. The priority should now be to exploit the potential of this path for research.

## References:

- Baltagi, B.H. (1995): *Econometric Analysis of Panel Data*. Chichester/New York/Brisbane/Toronto/Singapore.
- Bender, S. and Möller, J. (2010): Data from the Federal Employment Agency. [In this publication].
- Bundesministerium für Arbeit und Soziales (2008): *Lebenslagen in Deutschland – Der 3. Armuts- und Reichtumsbericht der Bundesregierung*. Bonn.
- Falk, M. and Steiner, V. (2000): *Expertise für die Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik im Bereich Erwerbstätigkeit und Einkommen*. Mannheim.
- Hauser, R. (2010): Household income, poverty, and wealth. [In this publication].
- Heinzel, A. (2006): *Volkszählung 2001 – Deutschland bereitet sich auf den registergestützten Zensus vor*. *Berliner Statistik* 7/ 2006, 321-328.
- Körner, T./Nimmergut, A./Nökel, J. and Rohloff, S. (2006): *Die Dauerstichprobe befragungsbereiter Haushalte*. *Wirtschaft und Statistik* 5/2006, 451-467.
- Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik (KVI) (Ed.) (2001): *Wege zu einer besseren informationellen Infrastruktur*. Baden-Baden.
- Robinson, W.S. (1950): *Ecological Correlations and the Behavior of Individuals*. *American Sociological Review* 15, 351-357.
- Rudolph, H. (2008): *Erwerbstätigkeit in Transferhaushalten – Arm trotz Arbeit*. *IAB Forum* 2/2008, 34-39.
- Schneider, H. and Wolf C. (2008): *Die Datenservicezentren als Teil der informationellen Infrastruktur*. In: Rolf, G./Zwick, M. and Wagner, G.G. (Eds): *Fortschritte der informationellen Infrastruktur in Deutschland*. *Festschrift für Johann Hahlen zum 65. Geburtstag und Hans Jörg Krupp zum 75. Geburtstag*. Baden-Baden.
- Statistische Ämter des Bundes und der Länder (2004): *Ergebnisse des Zensus 2004*. *Wirtschaft und Statistik* 8/2004, 813-833.
- Wagner, J. (2010): *Firm-Level Data*. [In this publication].

4.3 Interdisciplinary Longitudinal Surveys  
**Linking Individual Data to Organizational Data in  
Life-Course Analysis**

*Stefan Liebig*

Contact:

Stefan Liebig  
Bielefeld University  
Faculty of Sociology  
Postfach 10 01 31  
33501 Bielefeld  
Germany  
e-mail: stefan.liebig[at] uni-bielefeld.de

## Abstract

This paper starts with three fundamental insights from social science and economics: (1) that the conditions and consequences of individual behavior can only be studied empirically on the basis of longitudinal data, (2) that individual behavior is embedded in social contexts and social aggregates, and (3) that formal organizations – e.g., firms, schools, universities – are becoming more important for individual life courses. From this, it follows that social and economic research needs a data infrastructure which provides information on individuals over time and on the organizations those individuals are associated with. In the last nine years, there have been major efforts to provide scientific communities with linked individual-firm data in Germany. However, the available datasets comprise only limited information on individuals and organizations and provide no information on the household level. As the latter is becoming more important – e.g., in generating social inequalities – the existing data-stock should be complemented by longitudinal data linking individuals, their households, their firms, and other organizations they are members in. The recommendation is to enhance the German Socio-Economic Panel (SOEP, *Sozio-oekonomisches Panel*) with information from the firms the household members are presently employed in. Such a dataset can be useful for a wide range of social and economic research areas and would be unique on an international level.

Keywords: longitudinal data, surveys, linked employer-employee data, microdata, household data, life-course analysis, survey methodology

## 1. Research questions

One of the fundamental insights gained by the social and economic sciences is that empirically founded statements on the conditions and consequences of individual behavior or of social and economic change can only be formulated on the basis of longitudinal microdata. The observation of individuals, households, and other socio-economic units over long periods of time allows us to causally determine the reasons for social and economic stability and change. Moreover, socio-economic phenomena are particularly path-dependent. The opportunities and restrictions that individual or corporate actors face over their life courses – or more generally: over time – depend to a great extent on decisions and events earlier in time. The available individual and household-level datasets used in empirical social and economic research in Germany are capable of mirroring these path-dependencies.

But social and economic phenomena show another fundamental quality: they are embedded in social contexts and social aggregates (Granovetter 1985). Embeddedness means that actors are in most cases elements of a number of social aggregates. Their behavior is affected by these different

memberships and the structures and processes that take place within these aggregates, whether households, social networks, schools, firms, associations, regional areas, or nations. Longitudinal microdata for assessing the effects of these different social contexts on individual decisions and behavior are available at the level of households, geographic units, or – within comparative research – at the national level. However, recent labor market and educational research shows that there is another type of social aggregate that is crucial for an individual's economic or social situation and his or her life chances: institutions and organizations like schools, universities, firms, or establishments (Baron and Bielby 1980; Coleman 1993; Hamermesh 2008; Heckman 2001).

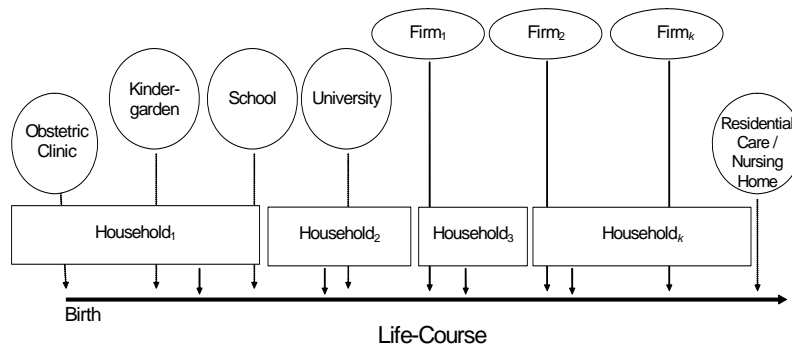
For many years, organizations have played a subordinate role in German research on social stratification, the labor market, and the education system (Allmendinger and Hinz 2002). With regard to firms and establishments, this was justified with reference to the dominance of the tariff system and the longstanding practice of macro-level regulation. Today, there exist a range of empirical studies showing a general trend towards increasing heterogeneity on the organizational level in Germany and suggesting that labor market and educational institutions are developing more and more differentiated internal structures and processes. One consequence of this development has been that the distribution of goods, jobs, and life opportunities is determined increasingly by the "internal logic of organizations." Some of the main effects of this on individual career paths and employment histories can be seen at the establishment level and firm level (Bender et al. 2000; DiPrete et al. 2001), on the level of wages (Kölling et al. 2005), in the duration of unemployment, and in qualification levels (Frederiksen et al. 2006), and even in the political attitudes of employees (Liebig and Krause 2007). Besides the classical variables such as number of employees (Heyman 2007), degree of unionization (Fitzenberger et al. 2007), and branch affiliation, a range of other important explanatory factors can be identified on the firm level and establishment level, such as a firm's age (Brixy et al. 2007), its socio-demographic structure (Krell and Sieben 2007), the magnitude of income disparities or mobility chances (Liebig and Krause 2007), and the particular form of work organization (Bellmann and Pahnke 2006).

The operative structures, processes, and strategies, as well as the business situations of employers are becoming increasingly important, and not only for employment revenues (Goedicke 2006; Lengfeld 2007). The variety of firm-specific operative time regimes, improvements in the compatibility between work and family, health promotion activities, and more flexible regulations governing working time and location (e.g., home workplaces) also affect an individual's social relations and his or her way of life in general (Düntgen and Diewald 2007).



As has already been outlined, organizations can control their members' access to jobs and goods. This is an assumption that takes on particular importance when analyzing durable structures of social inequality (Tilly 1998). The individual life course can also be understood as a sequence of different memberships in organizations (Figure 1). Individual life courses can thus be distinguished by the extent to which people succeed in joining organizations that offer better life chances. In this context, social stratification research tries to investigate whether this also results in path-dependencies, i.e., as people become members of advantageous or disadvantageous organizations, advantages and disadvantages are accumulated over the life course.

Figure 1: The individual life course and memberships in different types of organizations



## 2. Status quo

In order to empirically analyze the effects of the organizational level on individual career paths, the conditions and outcomes of employment, and different aspects of individual life courses, social and economic research requires adequate data linking personal and organizational information. Such matched organization-member datasets are available especially in the field of labor market research. These Linked Employer-Employee (LEE) datasets are characterized by a hierarchical multilevel structure, in which employees constitute the bottom level and the firms and/or enterprises constitute the upper level. The distinct feature of these LEE data is that they contain information about several – and in the optimal case, all – persons employed in a firm. In most cases, “process-produced” administrative data, on either the individual or the firm level, constitute the basis of analysis (Abowd and

Kramarz 1999). In contrast to other European and non-European countries, Germany recognized the potential of LEE data very late. This is why, in 2001, Martin Falk and Viktor Steiner concluded, in their advisory report to the 2001 KVI report: “The opportunities of matching firm and individual data were recognized much earlier in other countries. In certain areas, such as operative employment and income trends, German research is no longer competitive. In this domain, research is almost non-existent” (p. 8).

In the meantime, the data supply has been improved substantially, mainly because of the linked employer-employee dataset from the IAB (LIAB) (Alda et al. 2005) and the income and wage structure surveys conducted by the official statistical agencies (Stephan 2001), which are available in the Research Data Centers of the Federal Employment Agency at the Institute for Employment Research (IAB, *Institut für Arbeitsmarkt- und Berufsforschung* within the BA, *Bundesagentur für Arbeit*) and the Research Data Centers of the Federal Statistical Office and the Statistical Offices of the German *Länder*. Both data sources are “real” linked employer-employee datasets that offer information on all – or at least a sufficient number of – employees in each participating firm. Both datasets contain vast and diverse potential for analysis. The central difference is the degree of available firm information contained. The income and wage structure survey is a cross-sectional dataset; it only contains the basic parameters of the employment structure, sectoral affiliation, and degree of collective bargaining. Thus, it can be used primarily for the analysis of cross-sectional wage structures (especially after the inclusion of surveyed firms and sectors through changes in the legislation in January 2007). The LIAB, on the other hand, offers a broader base of information, ranging from detailed employment structures, the firm’s economic situation, professional training programs, to labor time regulations, payment systems, and special measures to improve compatibility between work and family. Although this focus indeed requires further development – e.g., with regard to the existing mobility regimes or the firm culture, which is quite important to organizational research – on the operative side, the LIAB offers a potential for analysis that exceeds the classic labor economic or sociological questions, all the more so because it displays longitudinal processes on the firm level and on the individual level. This central advantage is diminished, however, by the restricted supply of information on the employee side. Here, the LIAB shares one of the main weaknesses of the income and wage structure survey.

Both available LEE datasets in Germany are characterized by restricted access to information on individuals and households. This applies to central features of current employment relationships (the LIAB does not identify, e.g., temporary employment or the supply of temporary workers), to information on the economic situation of an individual, and even more so to household data, the family situation, social origins, social preferences and

personal characteristics, norm and value orientations, and political attitudes and membership in parties or other organizations. Since these topics are of central interest in empirical social and economic research, there is a strong need for a dataset that contains longitudinal information on the individual, household, and organizational level.

Against this background, an extension of the existing linked employer-employee data supply in Germany is desperately needed. This improvement needs to be promoted especially for the kind of information that goes beyond basic employment data. This can be achieved, for instance, by gathering information on family background, family and domestic situations, integration into social networks, as well as moral concepts and political attitudes. Improvements are also possible on the organizational side – the data catalog of the IAB Establishment Panel can, for instance, be expanded to include income and wage formation processes, elements of enterprise and firm culture, industrial relations, and the national or international competitive position of firms. Such a catalog of information can only be created on the basis of linked employer-employee surveys. In the present research, these kinds of data are produced using two different approaches:

### *2.1 Employer-first approach*

In the first step of this approach, which has also been pursued by official statistics in the framework of the income and wage structure survey or the WeLL<sup>1</sup> project by IAB and RWI (Bender et al. 2008), suitable firms are selected. Individual information is collected from a sample of employees working in these firms (either all employees or a partial sample) (see the 2000 National Employer Survey, Capelli 2001). The advantage here is that the existing multilevel data structure prevailing in common LEE datasets is still existent. One problem, however, is that such samples quite rapidly go beyond realistic limits. This happens if the information on the employees is not supplied by the firm itself but gathered by employee surveys. The coordination and implementation of such employee surveys in more than 100 or 200 firms is hardly practicable in the framework of normal research projects – even when the surveys are conducted by survey institutes. Accordingly, a recent project in Germany utilizing this approach concentrated on a single-digit number of firms (Brose et al 2006).

---

1 Further Training as a Part of Lifelong Learning (*Berufliche Weiterbildung als Bestandteil Lebenslangen Lernens*).

## 2.2 Employee-first approach

In the second approach to generating matched datasets, not firms or organizations, but persons (employees), constitute the point of departure. The individual data, which are gathered through personal interviews, are later complemented by firm-level data. This again can be done in three different ways (a technique that is already being used in research projects) (see Kmec 2003):

- (1) The information on the establishment or firm where the respondent of a population survey is employed are added using available commercial business datasets (in Germany: Creditreform or Hoppenstedt). Examples of this method are the New Worker Establishment Characteristics Database and the Decennial Employer-Employee Dataset. The problem of this approach is the limited scope of available firm information in the databases (e.g., number of employees, founding year, business volume). Although business databases can be used to assess an enterprise's liquidity or financial strength, they are less suitable for scientific questions.
- (2) The second way is to complement the personal information with data from official statistics for the appropriate establishment or firm. In the framework of a study conducted by the Max Planck Institute for Human Development in Berlin, for instance, researchers asked the respondents for their social security numbers. Afterwards, the individual data were linked to the IAB Establishment Panel (Reimer and Kuenster 2004). If the employer was included in the IAB Establishment Panel, the firm information was added to the individual data record. Obviously the problem here is that the share of employees in a population survey who are covered at the same time by the IAB Establishment Panel is expectedly small. Another possibility lies in using the IAB Establishment History Panel (Dundler et al. 2006), but in this case, the available employer information is much more restricted than in the IAB Establishment Panel.
- (3) Finally, one can use an individual or household survey to ask employees for the name and address of their employer, and can conduct a separate firm survey on the grounds of this information. The collected firm-level data can then be matched with the individual or household data. Examples of this approach are the Multi-City Study of Urban Inequality and most notably the National Organization Survey (NOS) from the years 1991 and 2002 (Kmec 2003). In the framework of the German General Social Survey (ALLBUS, *Allgemeine Bevölkerungsumfrage der Sozialwissenschaften*) of 1991 and 2002, all (1991) and, respectively, some (2002) of the currently employed were asked for the name and address of their workplace. Local business units in which people were gainfully

employed were the target units. On the basis of these entries, telephone interviews were conducted and postal questionnaires distributed. These data were matched to the individual data of the ALLBUS. The result is a linked employee-employer dataset (Kalleberg et al. 1996; Smith et al. 2004). In 1991, for a total of 51 percent of all cases (in 2002, 48 percent) the individual and firm-level data could successfully be matched. In contrast to the classical LEE data structure, this dataset does not possess a hierarchical structure. For one firm, the individual data are available for just one employee. Due to its cross-sectional character, this does not offer causal or longitudinal potentials for analysis. But through combined individual-firm surveys, it is possible to collect far more firm information than in a person-to-person interview, and the firm-level data, which are collected in combined surveys, are gathered independently from the interviewee's attitudes and perceptions (see Gupta et al. 2000).

In a current project underway at the University of Bielefeld, the design of the NOS study is being tested for its transferability to the German situation. For this purpose, all currently employed persons who are being surveyed in the ALLBUS 2008 (a nationwide reference survey) were asked for the name and addresses of their employers. Useful data is available for about 85 percent of those people who are employed in firms with more than six employees. On the basis of these data, a firm survey will be conducted in January 2009. The aim of this study is to assess the quality and methodological problems arising in connection to the generation of survey-based LEE datasets. Moreover, conclusions for future interview projects will be derived. As the willingness to participate in firm surveys has decreased constantly since the 1990s, another important task will be to find ways to maximize firm participation. A central problem of such a twofold survey-based approach is data protection. The respondents have to give permission for their firms to be contacted. Only then can individual and firm information be matched. A further problem is the re-identification of individuals and firms. However, the projects currently carried out by the official statistical agencies on the anonymization of firm and panel data already offer suitable tools that simplify data access – also for researchers.

### **3. Future developments**

Empirical research shows that there is an increasing variety of organizing work at the firm level in Germany which affects labor market processes, social stratification, and other socio-economic phenomena (e.g., work-life balance). From this follows an increasing demand for socio-economic data-

sets that identify linkages between individuals and organizations. Especially in the field of educational research, the interest in particular educational institutions will increase in the near future (Klieme 2008). The efficiency and the evaluation of activities will be measured according to their impact on the student's performance and his/her educational achievements. However, if no further household information is available, the linkages between organizational and individual data are not sufficient – especially with regard to the educational system.

The linkages between different data sources (e.g., Bender et al. 2007) offer the chance to broaden the scope of survey-based organizational data and to match them with information from other data sources. This reduces interview costs and allows the researcher to conduct firm surveys that are more strongly focused on a specific topic. As socio-economic research has recognized the need for longitudinal data and the embeddedness of individual behavior, it seems to be more important than ever before to collect longitudinal information on the individual and the household level.

#### **4. Recommendations**

Against this background the following recommendations can be made:

- (1) There is an increasing demand for linked data between individual, household, and organization information – especially with regard to the organization of the educational system and the workplace.
- (2) As the available datasets only offer limited information, household and individual surveys should be matched based on adequate organizational data. This can be achieved by matching data from official statistics or from separate surveys.
- (3) Linked individual/household and organizational datasets will be only feasible for socio-economic research if they contain longitudinal information.
- (4) The best solution to achieve an adequate data structure is to enrich the SOEP with separate firm surveys (e.g., of nursery schools, schools, workplaces of other household members) at five-year intervals. Respondents to SOEP should be asked for the names and addresses of these organizations, and based on this information, organizational surveys should be conducted to achieve a three-level hierarchical and longitudinal dataset. In this way longitudinal information would be made available on the individual, the household, and the organizational level. Such a dataset would be internationally unique and would offer novel

potential for analysis in a variety of disciplines (education, sociology, economics, psychology).

## References:

- Abowd, J.M. and Kramarz, F. (1999): The analysis of labor markets using matched employer-employee data. In: Ashenfelter, O. and Card, D. (Eds.): *Handbook of Labor Economics*. Amsterdam.
- Abowd, J.M./Haltiwanger, J. and Lane, J. (2004): Integrated Longitudinal Employer-Employee Data for the United States. *American Economic Review*, (94), 224-229.
- Alda, H./Bender, S. and Gartner, H. (2005): The linked employer-employee dataset of the IAB (LIAB). IAB Discussion Paper 6/2005.
- Allmendinger, J. and Hinz, Th. (2002): Perspektiven der Organisationssoziologie. In: Allmendinger, J. and Hinz, Th. (Eds.): *Organisationssoziologie*. Sonderheft der *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 42/2002. Opladen.
- Baron, J.N. and Bielby, W.T. (1980): Bringing the Firms Back, Stratification, Segmentation, and the Organization of Work. *American Sociological Review*, (45), 737-765.
- Bellmann, L. and Pahnke, A. (2006): Auswirkungen organisatorischen Wandels auf die betriebliche Arbeitsnachfrage. *Zeitschrift für Arbeitsmarktforschung*, (39), 201-233.
- Bender, S./Konietzka, D. and Sopp, P. (2000): Diskontinuität im Erwerbsverlauf und betrieblicher Kontext. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, (52), 475-499.
- Bender, S./Wagner, J. and Zwick, M. (2007): Kombinierte Firmendaten für Deutschland. FDZ Methodenreport 5/2007. Nuremberg.
- Bender, S./Fertig, M./Görlitz, K./Huber, M./Hummelsheim, S./Knerr, P./Schmucker, A. and Schröder, H. (2008): WeLL – Berufliche Weiterbildung als Bestandteil Lebenslangen Lernens. Projektbericht. FDZMethodenbericht 05/2008. Nuremberg.
- Brixy, U./Kohaut, S. and Schnabel, C. (2007): Do Newly Founded Firms Pay Lower Wages? First Evidence from Germany. *Small Business Economics*, (29), 161-171.
- Brose, H.G./Goedicke, A./Diewald, M. (2006): Beschäftigungsverhältnisse und sozialer Tausch. Wechselwirkungen zwischen Arbeitsangebot und Arbeitsnachfrage. Universität Duisburg-Essen/Universität Bielefeld. <http://www.uni-due.de/beata/>. [Last visited: 02/19/2010].
- Capelli, P. (2001): The National Employer Survey: Employer Data on Employment Practices. *Industrial Relations*, (40), 635-647.
- Coleman, J.S. (1993): The Rational Reconstruction of Society. *American Sociological Review*, (58), 1-15.
- Diewald, M./Brose, H.-G. and Goedicke, A. (2005): Flexicurity im Lebenslauf. Wechselwirkungen zwischen pluralen Lebensformen und betrieblichen Beschäftigungspolitiken. In: Kronauer, M. and Linne, G. (Eds.): *Flexicurity. Die suche nach Sicherheit in der Flexibilität*. Berlin.
- DiPrete, Th.A./Goux, D./Maurin, E. and Tahlin, M. (2001): Institutional Determinants of Employment Chances: The Structure of Unemployment in France and Sweden. *European Sociological Review*, (17), 233-254.



- Dundler, A./Stamm, M. and Adler, S. (2006): Das Betriebs-Historik-Panel. FDZDatenreport 3/2006.
- Düntgen, A. and Diewald, M. (2007): Auswirkungen der Flexibilisierung von Beschäftigung auf eine erste Elternschaft. In: Szydlík, M. (Ed.): Flexibilisierung. Folgen für Arbeit und Familie Wiesbaden.
- Falk, M. and Steiner, V. (2001): Expertise für die Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik im Bereich Erwerbstätigkeit und Einkommen. In: Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik (KVI) (Ed.): Wege zu einer besseren informationellen Infrastruktur. Baden-Baden.
- Fitzenberger, B./Kohn, K. and Lembcke, A.C. (2007): Union Density, Collective Bargaining, and Individual Coverage: The Anatomy of Union Wage Effects. Mannheim.
- Frederiksen, A./Ibsen, R./Rosholm, M. and Westergaard-Nielsen, N. (2006): Labor Market Signalling and Unemployment Duration: An Empirical Analysis Using Employer-Employee Data. IZA Discussion Paper No. 2132.
- Goedicke, A. (2006): Organisationsmodelle in der Sozialstrukturanalyse: Der Einfluss von Betrieben auf Erwerbsverläufe. Berliner Journal für Soziologie, (16), 503-524.
- Granovetter, M. (1985): Economic Action and Social Structure: The Problem of Embeddedness. American Journal of Sociology, (91), 481-510.
- Gupta, N./Shaw, J.D. and Delery, J.E. (2000): Correlates of response outcomes among organizational key informants. Organizational Research Methods, (3), 323-347.
- Hamermesh, D. (2008): Fun with matched firm-employee data: Progress and road maps. Labor Economics, (15), 663-673.
- Heckman, J.J. (2001): Micro Data, heterogeneity, and the evaluation of public policy: Nobel lecture. Journal of Political Economy, (109), 673-748.
- Heyman, F. (2007): Firm Size or Firm Age? The Effect on Wages Using Matched Employer-Employee Data. Labour, (21), 237-263.
- Kalleberg, A.L./Knoke, D./Marsden, P.V. and Spaeth, J.L. (Eds.) (1996): Organizations in America. Thousand Oaks.
- Klieme, E. (2008): Längsschnittuntersuchungen zur Schulentwicklung. Vortrag auf der 4. Konferenz für Sozial- und Wirtschaftsdaten. [http://www.ratswd.de/kswd/download/4\\_KSWD\\_Klieme.pdf](http://www.ratswd.de/kswd/download/4_KSWD_Klieme.pdf). [Last visited 09/20/2008].
- Kmec, J.A. (2003): Collecting and Using Employer-Worker Matched Data. Sociological Focus, (36), 81-96.
- Kölling, A./Schnabel, C. and Wagner, J. (2005): Establishment Age and Wages: Evidence from German Linked Employer-Employee Data. In: Bellmann, L./Hübner, O./Meyer, W. and Stephan, G. (Eds.): Institutionen, Löhne und Beschäftigung, Beiträge zur Arbeitsmarkt- und Berufsforschung Nr. 294. Nuremberg.
- Krell, G. and Sieben, B. (2007): Diversity Management und Personalforschung. In: Krell, G./Riedmüller, B. and Sieben, B. (Eds.): Diversity Studies – Grundlagen und disziplinäre Ansätze. Frankfurt am Main.
- Lengfeld, H. (2007): Organisierte Ungleichheit. Wie Organisationen Lebenschancen beeinflussen. Wiesbaden.

- Liebig, S. and Krause, A. (2007): Arbeitsorganisationen als Kontexte der Einstellungsbildung. In: Hummell, H.J. (Ed.): Die Analyse von Gesellschaften, Organisationen und Individuen in ihrem Zusammenhang – Theoretische und methodische Herausforderungen. Tagungsberichte Band 13. Bonn.
- Reimer, M. and Kuenster, R. (2004): Linking Job Episodes from Retrospective Surveys and Social Security Data: Specific Challenges, Feasibility and Quality of Outcome. Arbeitspapier Nr. 8 des Projekts Ausbildungs- und Berufsverläufe der Geburtskohorten 1964 und 1971 in Westdeutschland. Berlin.
- Schnabel, C. (2006): Verbetrieblischung der Lohnfindung und der Festlegung von Arbeitsbedingungen. Arbeitspapier 118. Düsseldorf.
- Smith, T.W./Kalleberg, A.L. and Marsden, P.V. (2004): National Organization Survey (NOS), 2002. Ann Arbor.
- Tilly, Ch. (1998): Durable Inequality. Berkeley.

## 4.4 Organizational Data

*Stefan Liebig*

Contact:

Stefan Liebig  
Bielefeld University  
Faculty of Sociology  
Postfach 10 01 31  
33501 Bielefeld  
Germany  
e-mail: stefan.liebig[at] uni-bielefeld.de

## **Abstract**

Organizational data describe central characteristics of organizations, their internal structures and processes as well as their behavior as corporate actors in different social and economic contexts. Firm and enterprise data are the most frequently used type of organizational data, but there is also a growing interest in data on schools, universities, and hospitals in the economic and social science research. In the last several years, there has been a substantial improvement in the accessibility and scientific usability of organizational data from official statistics. However, non-official organizational data produced within publicly funded research projects are practically impossible to obtain for secondary analyses. There is no documentation of the existing stock of non-official organizational data, and the methodological standards used for organizational research in Germany are low compared to the standards of international research. Against this background, it is recommended that efforts be focused on documenting and archiving the existing non-official organizational data for secondary analyses and on establishing higher methodological standards within this research field.

Keywords: firms, organizations, methods of organizational research, microdata, secondary analysis

## **1. Research questions**

The most common form of organizational data used in economic and social research relates to firms as local production units, in which goods and services are produced, and to enterprises as the legal units of the private and public sector. The data describe central characteristics of these organizations, their internal structures and processes, as well as their behavior as corporate actors in different contexts. Besides these kinds of “classic” firm-level data, data referring to organizations within the educational system (nursery schools, schools, and universities) have recently also attracted attention in Germany (Klieme 2008). This interest has arisen in the context of an increasing awareness that the structures and processes existing on the school level – demographic composition of the school, quality of the cooperation among staff members – are important for individual educational success. Furthermore, an ongoing differentiation is being observed on the level of individual organizations within the German educational system, making it more important, for example, which university a person graduated from.

Each of the different disciplines focus on describing and explaining different structures and processes of organizations and their actions. The organizational research (business administration, sociology, psychology) is preoccupied with the structural characteristics of firms (degree of centrali-

zation, formalization, and standardization), the internal forms of organization of work, the design and practice of operational staff policy, the industrial relations or the reasons for growth and shrinking of firms. Moreover, organizational-level data offer the possibility to evaluate the effects of policy measures. In this case, the structural features of organizations and their behavior are the objects that are to be explained. The question here is how organizations react to changes in the legal, economic or social surroundings; in other words, which effects specific changes in the social, economic, and legal environment have on organizations. Vice versa, firm-level data can additionally be used for the explanation of other issues such as macro economic developments, job market dynamics, educational participation or the reproduction of social inequality. In this respect, external consequences of organizational behavior or their internal structures and their changes stand in the center of interest. In this context, organizations represent micro-level units that help to explain macro-level phenomena. Correspondingly, economic researchers have defined firm-level data thus far as microdata. Important questions are, for instance, the consequences of operative employment trends, apprenticeship and advanced training, productivity and investment in the different areas of interest to economic policy (see Wagner 2010). Topics of organization-centered education research are the relevant surroundings and arrangements of educational organizations, their composition with respect to their personnel (teachers) and clients (children, pupils, students), and the resulting effects on education performance, education participation, and social inequalities (see Klieme 2008; Garmoran et al. 2006; OECD 2005).

To answer these questions, information about organizations can be collected in two different ways: first, through primary data collection using reactive and non-reactive research methods. This can take the form of interviewing (reactive methods), in which information about the organization is gathered by persons inside or outside the organization using standardized (surveys) or non-standardized (case studies) questionnaires (Bryman 2000, Stablein 1999). In the framework of non-reactive methods, data on firm structures, processes, and behavior can be collected by making use of documents provided by the organization itself or archives in which data about organizations are stored. Second, organizational data can be collected by gathering information that accompany administrative processes. These “process-produced” administrative data arise either within organizations, e.g., in personnel administrations (see Brüderl et al. 1993), or outside organizations, e.g., in social security and tax administrations (see Wagner 2010). Currently, the most frequently used organizational data in Germany are survey data, qualitative case studies, and “process-produced” administrative data.

One of the major findings of organizational research is that organizations are not well coordinated units that follow strict and coherent bureaucratic principles (Hannan and Freeman 1977; Sørensen 2007). A number of organizational scholars hold the view that the complexity and diversity of organizations can solely be represented adequately on the basis of case studies with the aid of qualitative survey methods. As a result, the organizational research has given rise to a multitude of qualitative case studies. The use of qualitative methods and the concentration on case studies has long been a distinctive feature of the German organizational research by international comparison (see Grunow 1995). Case studies do in fact have an important heuristic function in the research process, yet they entail some major problems: due to the small number of cases, they cannot provide any generalized statements; inter-subjective validation of the findings is impossible; and they can have only limited significance for social and economic research, which is more oriented towards explaining and predicting phenomena (see Hauptmanns and Rogalski 1992). Accordingly, there is a need for quantitative firm-level data based on standardized survey methods that allow for utilization of econometric methods. At the same time, such datasets need to have a sufficiently large sample size: only on this basis can researchers conduct analyses on the level of economic sectors, sub-sectors, and regional units. Moreover, there is a need to adequately describe changes over time and to scrutinize causal explanatory models. Such questions can only be answered by analyzing longitudinal data. This is why panel data have taken a central role in organization-related social and economic research since the 1990s (Heckman 2001; Wagner 2008). Regarding the collection and provision of quantitative data on economic organizations in Germany, the situation since the 1980s has been as follows: most of the data outside official statistics were acquired by analyzing cross-sectional studies restricted to smaller, individual economic sectors, distinct types of enterprises, and single regions.<sup>1</sup> Only in the 1990s were a number of larger longitudinal firm-level datasets generated (e.g., the *NIFA-Panel*), which were then expanded to intersectoral and national scales (IAB Establishment Panel). Since that point, firm-level data from official statistics have been gradually made accessible to researchers (KVI 2001). However, only since 2001, with the creation of Research Data Centers, have these data found real applications in research. These conditions have not yet been achieved for data on organizations within the educational system (Stanat 2008).

Organizational data are only valuable for scientific purposes if data production is guided by methodological standards, and if the resulting fin-

---

1 The only exceptions are the surveys within the framework of official statistics, the process-produced data of the Establishment History Panel (*Betriebs-Historik-Panel*) of the former Federal Employment Agency (*Bundesanstalt für Arbeit*), as well as a few individual studies, such as the Ifo Business Survey.

dings can be reconstructed by other researchers and assessed on the basis of the data. The latter is only possible if data are made available in a broadly useable form. When it comes to methodological standards, most of the organizational research is concerned only with the quality of data analyses. But there is also the problem of data collection, which raises questions of survey methodology even in organizational research: the quality of data depends not only on the sample or the sampling procedures used, but also on the validity and reliability of the measures. In this respect one must ask: who in the organization provides the information on what basis? Does a question measure the same phenomenon in different sectors, sub-sectors, or firms?

The plea for making organizational data available is confronted by a fundamental problem of organizational research: with very little information, it is relatively easy to re-identify the firms and enterprises from which the data was collected (see Gottschalk 2002). Data collected on the basis of the compulsory duty-of-disclosure of official statistics, however, can only be made accessible if re-identification is impossible (*BStatG §16*)<sup>2</sup>. This does not apply to firm-level data, which are generated through voluntary participation. Nevertheless, anonymity is needed in order to convince firms to participate. The protection of participants' confidentiality prevents non-anonymous use of the data by a third party. However, in the last years, new methods and techniques have been developed that allow anonymization of firm-level data without diminishing its worth for scientific research (Drechsler et al 2007; Rosemann 2006; Wagner 2010).

## 2. Status quo

The organizational data used in social and economic research can be subdivided into three different groups: (1) data from commercial providers, (2) survey or "process-produced" administrative data from official statistics, and (3) data collected by research institutions or individual researchers.

(1) Currently, commercial firm-level data are available solely as enterprise data. The two most important databanks in Germany are the *Hoppenstedt-Firmendatenbank* and the *Creditreform-Firmenprofile*. Both contain a limited number of details (e.g., form of organization, sales figures, number of employees over the past years, contacts on top management level) about the enterprises as legal units. Therefore, they offer no information about the firms in terms of local units. Both suppliers exclude specific enterprise groups. *Creditreform*, for instance, rules out certain legal forms, *Hoppenstedt* excludes enterprises with an annual turnover below 1 million

---

2 The citations to German legal sources have been left in German to guarantee accuracy.



euros or with less than 20 employees. The information is based on entries in commercial registries or on the suppliers' own research, in which case the entry is voluntary. An additional commercial dataset is *LexisNexis*, which collects information about enterprises thereby using different sources – *Hoppenstedt*, *Bundesanzeiger*, commercial registries, and press releases.

Besides the restricted information base for many scientific and applied questions, the problems of these data are that (1) data collection is not documented and not transparent, and (2) the enterprises listed do not necessarily constitute the respective universe. The access to data for researchers, however, is quite good, as the suppliers have specific and less expensive offers for scientific purposes. In addition, databanks are made available as a standard part of many universities' research resources. Moreover, there exists a range of other national, international, and comparative datasets on firm policies – also related to personnel policies – that are collected by private companies and consulting firms. They offer a broad range of information, but are in most of the cases not available for scientific use.

(2) Organizational data from official statistics are the firm and enterprise data collected by the Federal Statistical Office and the Statistical Offices of the German *Länder*, the Deutsche Bundesbank, and the Federal Employment Agency (BA, *Bundesagentur für Arbeit*). The data from the statistical offices are collected by legal order via surveys or are the result of administrative processes – e.g., reporting to the social security system or tax administration. The data are provided by the Research Data Centers of the Federal Statistical Office and the Statistical Offices of the German *Länder* via (1) Public Use Files with very restricted information, (2) Scientific Use Files, which are de facto anonymized,<sup>3</sup> (3) teleprocessing, and (4) on-site usage of the original data within the centers. The surveys are conducted separately for each economic sector, thereby producing different kinds of datasets that range from monthly figures and total surveys up to annual sample surveys (see: Brandt et al. 2007; Kaiser and Wagner 2008). Data on individual economic sectors (producing industry, trade, hotel industry, service sector) are accessible in the same way as other cross-sector surveys on wages and tax statistics.

Besides the data from the statistical offices, the BA offers two datasets: one of them is the Establishment Panel of the Institute for Employment Research (IAB, *Institut für Arbeitsmarkt- und Berufsforschung*), which is an annual, voluntary survey that offers information on 16,000 firms since the year 1993 (in East Germany: 1996). Of all the available official statistics, the IAB Establishment Panel offers the broadest scope of information. Quite recently, the Establishment History Panel has also been made available. As a “process-produced” administrative dataset, it aggregates information on

---

3 De facto anonymous data contain information that can only be traced back to the participant with time consuming and cost-intensive effort.

employees covered by the social security system in the period between 1975 and 2005 to the firm level, and creates a data stock with information on 1.5 to 2.5 million firms. Both datasets are accessible in the Research Data Center of the BA at the IAB for on-site use, and the Establishment Panel is also available via teleprocessing.

Due to the firms' legal duty of disclosure (which, however, does not apply to the IAB Establishment Panel), official organizational data are characterized by high participation rates – with even sensitive questions being answered thoroughly – and large sample sizes. This is why differentiated analyses are possible, even in small regional units. As most of the available microdata on establishment and firms are longitudinal panel data, it is possible to analyze processes of change and to test causal explanatory models (Brandt et al. 2007).

Data documentation and data access have been considerably improved in the last few years. In this context, the research project “De facto anonymization of business microdata” (Lenz et al. 2006) has played a decisive role here, by developing solutions to the anonymization problem. The follow-up project “business panel data and de facto anonymization“ sets the ground for the expansion of available data (Scientific Use Files) through longitudinal panel data.

A central problem with organizational data from official statistics is that they only capture a small amount of information – mainly business and personnel statistics – which are useful only for specific fields of research in economics and the social sciences. Moreover, some of the surveys are restricted to single economic sectors. One possibility to resolve this problem is the inter-linkage of different datasets via the business register (which has been officially permitted since 2005). This is currently being investigated in the project Official Firm Data for Germany (*AFiD*, Konold 2007). The project Combined Firm Data for Germany (*KombiFiD*, Bender et al. 2007) even goes a step further by working on the linkage of data from the Federal Statistical Office, the Statistical Offices of the German *Länder*, and the Federal Employment Agency – which, however, is not yet legally permitted (see Wagner 2010). One central problem that has not yet been addressed by the ongoing projects is the quality of data collection (e.g., the problem of measurement error).

Data on organizations in the educational system – e.g., childcare facilities, schools, universities, as well as advanced training facilities – are collected by the Federal Statistical Office and the Statistical Offices of the German *Länder*, and are available on the micro-level at the respective Research Data Centers.

(3) Organizational data collected by research organizations or individual researchers generally have a smaller sample size. Since they are more strongly oriented towards substantive research questions, they contain more

information than the official data. In this regard, they offer a necessary extension of the official data sources, and they build the basis for organization-related research focusing more on the description and explanation of an organizational strategy, internal processes, industrial relations, etc. The relevant literature in this discipline shows that there exist a multitude of quantitative and qualitative (and in most cases cross-sectional) organizational studies, financed by public and private research organizations (e.g., DFG,<sup>4</sup> BMBF,<sup>5</sup> *VolkswagenStiftung*, *Hans-Böckler-Stiftung*, *Thyssen-Stiftung*). However, in contrast to the official statistics, the existing data stock is not documented and the data are not available for secondary analyses.<sup>6</sup> Exemptions are only the data from Ifo Institute for Economic Research (Ifo Business Survey, Becker and Wohlrabe 2008), Centre for European Economic Research (ZEW, *Zentrum für Europäische Wirtschaftsforschung*) in Mannheim (datasets on innovation, business trends and the middle classes), the NIFA-Panel (Widmaier 2000), as well as the longitudinal firm-level data that have been collected since 2001 in the framework of *Sonderforschungsbereich 580* (SFB580-A2 manager survey and SFB580-B2 establishment panel, Krause and Martens 2008). Currently, only the NIFA-Panel is incorporated into the data catalog of the GESIS<sup>7</sup> Data Archive, Cologne. German firm and establishment data on working time and work-life balance are available as part of an international comparative survey conducted by the European Foundation for the Improvement of Living and Working Conditions (Eurofound). The data are archived and accessible at the Economic and Social Data Service (ESDS), UK.

One of the main reasons for the inadequate availability of organizational data for secondary analysis might be the problem of re-identification. This is especially true for qualitative organizational data. But it seems that researchers within the field of organizations have not yet noticed the progress regarding anonymization methods of firm-level data made by the abovementioned projects.<sup>8</sup>

Closely connected to restricted usage and lack of data documentation are central methodological deficits, which are visible in many research articles dealing with organizational data. First, in most of the cases publications relying on organizational data do not offer any methodological explanations regarding sample quality or data collection. In addition, the organizational data lack a well-documented, standard methodological set of measurement

4 German Research Foundation (*Deutsche Forschungsgemeinschaft*).

5 Federal Ministry for Education and Research (*Bundesministerium für Bildung und Forschung*).

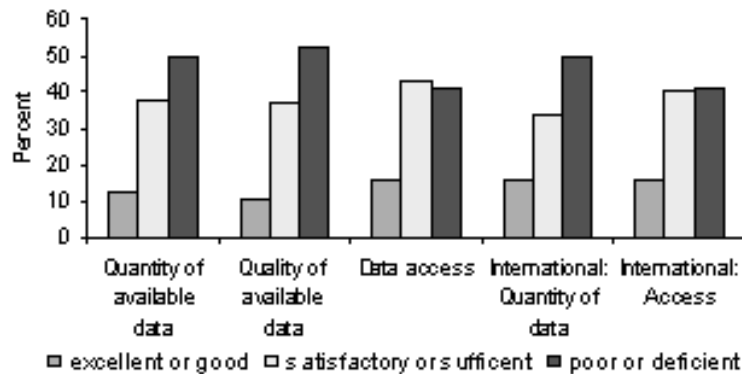
6 According to the author's own research, even the funding institutions do not have detailed information about previously collected data and their availability.

7 Leibniz Institute for the Social Sciences (*Leibniz-Institut für Sozialwissenschaften*).

8 The qualitative interview data that were collected in *Sonderforschungsbereich 580* are an exception.

instruments as it is common in survey research (see Statistisches Bundesamt 2004; Glöckner-Rist 2007). A strong need therefore exists to establish internationally comparable methodological standards and also for more research on organizational survey methodology, which concentrates on data collection methods, measurement errors, sampling, and unit- or item-non-response problems.

Figure 1: Results from an Expert Online Survey in Germany 2008.<sup>9</sup>



Notes: Relative Frequencies, Online Survey August/September 2008, N = 40.

The deficits in supply, methodological quality, and access – also in international comparison – have been articulated in an online survey among organizational researchers, which was conducted by the author and Alexia Meyermann in summer 2008. In this online survey, 50 percent of the participants assessed the data supply and the quality of content and methods as inadequate or insufficient. Forty percent criticized data access. A similar pattern can be found when looking at the numbers of participants who appraised the quantity of data and the access in international comparison. Overall, the researchers surveyed called for an improvement of the research situation. 74 percent of the researchers were prepared to make their data

<sup>9</sup> The online survey was conducted by Stefan Liebig and Alexia Meyermann. A call for participation was sent via mailing lists to different sections of the German Sociological Association (Sociology of Work and Industrial Sociology, Economic Sociology), to the research group on empirical personnel and organizational research (*Arbeitskreis Empirische Personal- und Organisationsforschung*) and to the German Industrial Relations Association (GIRA) in August/September 2008. The survey homepage was visited by 121 people, of whom 40 completed the survey.

available for secondary analyses, but had not delivered them to a data archive so far due to the anonymization problem.

### **3. Future development**

The future development of organizational research is characterized by an increasing demand for international comparative and longitudinal studies. This will be the only way to identify causal effects of the organizational level on macro-phenomena and vice versa. Especially under the conditions of an ongoing process of globalization, we will need this kind of data to study the relationship between macro-level economic processes and the behavior of organizations as corporate actors and their internal structures. Furthermore, European unification makes it necessary to broaden the narrow national perspective of organizational research. As the institutional conditions at the member state level are adjusted, we will need to investigate the accompanying restrictions and challenges for different organizations. For such international comparative and longitudinal organizational research, the implementation of methodological standards is necessary. At the same time, a broad base of information is decisive for the individual disciplines and for the evaluation of different policy measures. In the face of demographic changes and the debates on the work-life balance, not only personnel and business figures will be of interest but also the strategies and programs that are not covered by official statistics. Given the much higher obstacles to international comparative and standardized surveys, this will require the collection of non-official, comparative longitudinal organizational data. At the same time, the differentiation of the German educational system – which affects all levels – will increase the demand for a more detailed description of the internal structures and processes of nursery schools, schools, and universities (Klieme 2008).

All in all, the present progress in the field of official organizational data is positive. The improvements in data supply and data access are leading in the right direction. Further linkages among individual datasets and improved access – e.g., through remote access – have already been examined in different projects (see Wagner 2010). Furthermore, it will be essential not only to improve linkages within the official statistics, but to apply this same approach to the publicly and privately financed organizational surveys as well. While the official statistics offer exact longitudinal information on “hard” personnel and business figures, the advantage of non-official surveys can be seen in their thematic amplitude. Non-official organizational surveys can benefit from such linkages, since they do not bear the burden of data collection but can instead focus on specific research questions; initial efforts

in this field are already underway (for more on this issue, see Reimer and Künster 2004).

However, the enhancement of the organizational data infrastructure in Germany will also lead to another, perhaps less obvious problem: for the production and use of organizational data, specific competencies are required that are only taught in academia to very low degree at present. Although organizational research also relies on the methods and techniques of general empirical social research, standardized organizational surveys pose specific challenges with regard to methodological issues. These include questions of how to draw the sample, develop the sampling instruments, and collect the data, and require knowledge of the respective statistical techniques. This implies that data producers should increasingly offer CAMPUS-Files for scientific education. Moreover, methodological training in organizational research should be professionalized and intensified.

#### **4. Recommendations**

Against this background, the following recommendations for improving the existing infrastructure of organizational data can be given:

- (1) Documentation of existing non-official organizational data should be made easily accessible to interested researchers and enriched with detailed methodological information – at least including the publicly-funded data of the DFG, the BMBF, the Max Planck Society, and the Leibniz Association.
- (2) Data producers from universities and publicly financed research institutes should be obliged to make the data they collect available for research. As with the social data on persons and households, these data should be centrally archived. This should be done not only for quantitative data but also for qualitative organizational data (e.g., it is currently done in Sonderforschungsbereich 580) as is the case within the Economic and Social Data Service (ESDS), UK.
- (3) Research on organizational survey methods is urgently needed, as is enhanced academic training within the field of survey methods for organizational research.
- (4) A network of projects should be established that deals with the implications of data protection laws, practical solutions to the linkage of official and non-official organizational data, and the promotion of analogous policy measures.

- (5) The useable official statistics on organizational microdata should be enriched to include data on the educational system.
- (6) In addition to existing Research Data Centers, a specific Data Service Center on firm and organizational data should be established. The task of such a center should not only be to document the existing data on organizations in Germany and archiving the data from non-official producers but also to offer expertise and service for researchers who are planning organizational studies and want to provide their data for secondary analyses. Such a center should work on developing solutions for the anonymization of quantitative and qualitative organizational survey data and boosting the establishment of and adherence to methodological standards in order to improve the quality of organizational data in Germany. Only by creating a center that is responsible for documenting, archiving, and providing a broad range of methodological services, can the gap be closed between German organizational research and the international standards and infrastructures that exist in other countries (e.g., the Data Archive in UK).

## References:

- Becker, S.O. and Wohlrabe, K. (2008): Micro data at the IfO Institute for economic research – The “IfO Business Survey” usage and access. *Schmollers Jahrbuch* 127, 307-319.
- Bender, St./Fertig, M./Görlitz, K./Huber, M./Hummelsheim, St./Knerr, P./Schmucker, A. and Schröder, H. (2008): WeLL - Berufliche Weiterbildung als Bestandteil Lebenslangen Lernens: Projektbericht. FDZMethodenbericht 05/2008. Nuremberg.
- Bender, St./Wagner, J. and Zwick, M. (2007): Kombinierte Firmendaten für Deutschland. FDZ Methodenreport 5/2007. Nuremberg.
- Brandt, M./Oberschaltseik, D. and Pohl, R. (2007): Neue Datenangebote in den Forschungsdatenzentren Betriebs- und Unternehmensdaten im Längsschnitt. FDZ-Arbeitspapier 23. Wiesbaden.
- Brüderl, J./Preisendörfer, P. and Ziegler, R. (1993): Upward mobility in organizations. The effects of hierarchy and opportunity structure. *European Sociological Review*, (9), 173-188.
- Bryman, A. (2000): *Research Methods and Organization Studies*. London.
- Drechsler, J./Dundler, A./Bender, St./Rässler, S. and Zwick, Th. (2007): A New Approach for Disclosure Control in the IAB Establishment Panel. Multiple Imputation for a Better Data Access. IAB-DiscussionPaper 11/2007. Nuremberg.
- Gamoran, A./Secada, W. and Marret, C. (2006): The Organizational Context of Teaching and Learning. In: Hallinan, M.T. (Ed.): *Handbook of the Sociology of Education*.
- Glöckner-Rist, A. (Ed.) (2007): *ZUMA-Informationssystem. Elektronisches Handbuch sozialwissenschaftlicher Erhebungsinstrumente*. Version 11.00. Bonn.
- Gottschalk, S. (2002): Anonymisierung von Unternehmensdaten: Ein Überblick und beispielhafte Darstellung anhand des Mannheimer Innovationspanels. *ZEW Discussion Papers* 02-23.
- Grunow, D. (1995): Research Design in Organization Studies. *Organization Science*, (6), 93-103.
- Hannan, M. T., and Freeman, J. (1977). The population ecology of organizations. *American Journal of Sociology*, (82), 929-964.
- Hauptmanns, P. and Rogalski, W. (1992): Fallstudien in der Industriesoziologie – Zur Kritik der vorherrschenden Methode sozialwissenschaftlicher Technikforschung. In: Lehner, F. and Schmid, J. (Ed.): *Technik – Arbeit – Betrieb – Gesellschaft. Neue Informationstechnologien und flexible Arbeitssysteme* Bd. 1. Opladen.
- Heckman, J.J. (2001): Micro Data, heterogeneity, and the evaluation of public policy: Nobel lecture. *Journal of Political Economy*, (109), 673-748.  
[http://www.ratswd.de/kswd/download/4\\_KSWD\\_Klieme.pdf](http://www.ratswd.de/kswd/download/4_KSWD_Klieme.pdf). [Last visited 09/20/2008].
- Kaiser, U. and Wagner, J. (2008): Neue Möglichkeiten zur Nutzung vertraulicher amtlicher Personen- und Firmendaten. *Perspektiven der Wirtschaftspolitik* 9, 329-349.
- Klieme, E. (2008): Längsschnittuntersuchungen zur Schulentwicklung. Vortrag auf der 4. Konferenz für Sozial- und Wirtschaftsdaten.



- Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik (KVI) (Ed.) (2001): Wege zu einer besseren informationellen Infrastruktur. Baden-Baden.
- Konold, M. (2007): New possibilities for economic research through integration of establishment-level panel data of German official statistics. *Schmollers Jahrbuch* 127, 321-334.
- Konsortium Bildungsberichterstattung (2006): Zur langfristigen Sicherstellung der Datenbasis für die Bildungsberichterstattung. Frankfurt am Main.
- Krause, I. and Martens, B. (2008): Wann lohnt sich die Akquisition nicht mehr? Antworten auf der Basis von Erfahrungen mit telefonischen Expertenbefragungen im Paneldesign. In: Martens, B. and Ritter, Th. (Ed.): *Eliten am Telefon*. Baden-Baden.
- Lenz, R./Rosemann, M./Vorgrimler, D. and Sturm, R. (2006): Anonymising business micro data – results of a German project. *Schmollers Jahrbuch* 126, 635-651.
- OECD (2005): Chapter 3: The relative impact of school climate, school policies and school resources on quality and equity. In: OECD (Ed.): *School factors related to quality and equity. Results from PISA 2000*.
- Reimer, M. and Künster, R. (2004): Linking Job Episodes from Retrospective Surveys and Social Security Data: Specific Challenges, Feasibility and Quality of Outcome. Arbeitspapier Nr. 8 des Projekts Ausbildungs- und Berufsverläufe der Geburtskohorten 1964 und 1971 in Westdeutschland. Berlin.
- Rosemann, M. (2006): Auswirkungen datenverändernder Anonymisierungsverfahren auf die Analyse von Mikrodaten. Tübingen.
- Sørensen, J.B. (2007): Organizational diversity, labor markets, and wage inequality. *American Behavioral Scientist*, (50), 659-676.
- Stablein, R. (1999): Data in Organization Studies. In: Clegg, St.R. and Hardy, C. (Ed.): *Studying Organization. Theory & Method*, 255-271.
- Stanat, P. (2008): Daten der Bildungsforschung: Aktueller Stand und neue Entwicklungen. Vortrag auf der 4. Konferenz für Sozial- und Wirtschaftsdaten am 19. Juni in Wiesbaden.
- Statistisches Bundesamt (2004): *Demographische Standards*. Wiesbaden.
- Wagner, J. (2000): Firm Panel Data from German Official Statistics. *Schmollers Jahrbuch* 120, 143-150.
- Wagner, J. (2008): Improvements and Future Challenges for the Research Infrastructure in the Field Firm Level Data. Working Paper Series in Economics No. 88. Universität Lüneburg.
- Wagner, J. (2010): Firm Level Data. [In this publication].
- Widmaier, U. (2000): Das NIFA-Panel und der deutsche Maschinen- und Anlagenbau. In: Widmaier, U. (Ed.): *Der deutsche Maschinenbau in den neunziger Jahren. Kontinuität und Wandel einer Branche*. Frankfurt am Main.
- Zwick, M. (2007): CAMPUS Files. Free public use files for teaching purposes. *Schmollers Jahrbuch* 127, 655-668.



## 4.5 Firm-Level Data

*Joachim Wagner*

Contact:

Joachim Wagner  
Leuphana University Lüneburg  
Institute of Economics  
PO Box 2440  
21314 Lüneburg  
Germany  
email: [wagner\[at\]leuphana.de](mailto:wagner[at]leuphana.de)

## Abstract

This article discusses the use of enterprise- and establishment-level data from official statistics to document stylized facts, motivate assumptions used in formal theoretical models, test hypotheses derived from theoretical models, and evaluate policy measures. It shows how these data can be accessed by researchers in Germany today and reports on recent developments that will offer new and improved datasets that combine data collected in separate surveys and by different agencies. The paper makes three recommendations for future developments in this area: (1) change the law to make the combination of data collected by different producers easier, (2) combine firm-level data across national borders and make these data available for researchers, and (3) find ways to enable researchers in Germany to work with confidential firm-level data via remote access 24 hours a day and 365 days per year.

Keywords: firm-level data, Germany, FiDASt, KombiFiD, AFiD

## 1. What are firm-level data?

Firm-level data are data collected at, or related to local production units (establishments) or legal units (enterprises). The technical term used to describe this kind of data in official statistics is *wirtschaftsstatistische Einzeldaten*, or microdata for production units. This kind of data can either be collected in a survey (administered by a statistical office or by other institutions such as an opinion research institute or by a researcher at a university), or produced during a process that is related to administrative issues (for example, collection of taxes on sales or reporting to the social security system), resulting in what is named process-produced data.<sup>1</sup>

Usually, firm-level data are confidential – either by law (if they are collected in surveys from official statistics or are the outcome of administrative processes) or by an agreement between the (private, non-governmental) collector of the data and the firms that delivered the data. The reasons for confidentiality are manifold, including the fact that information delivered by firms that are required to report to surveys administered by official statistics has to be protected against competitors, and also that firms usually are only willing to respond to a survey voluntarily if they can be sure that any information considered to be “sensitive” will not be disseminated.

Confidentiality of firm-level data is a crucial issue for researchers who want to use microdata for production units in scientific studies. Although researchers are not at all interested in any of the establishments or enterprises

---

1 For a discussion of other organizational data (for example, data for organizations within the educational system) and publicly funded non-official organizational data collected by researchers, see the contribution by Stefan Liebig (2009).

per se, they need to access the data at the micro-level to perform their statistical analyses and econometric estimations and thereby to uncover patterns of firm behavior and test theoretical hypotheses. This paper discusses issues related to the use of confidential firm-level data by independent researchers (i.e., those who are not working for the data producers). It begins with a review of what firm-level data are good for (in section 2), who produces firm-level data in Germany, and how researchers can gain access to these data today (in section 3). In section 4, new and ongoing developments that are currently leading to new products are discussed – new types of firm-level data will considerably enhance the research potential available to researchers in the near future. Section 5 concludes with a wish list.

## 2. What are firm-level data good for?

Researchers use firm-level data in a wide range of areas in economics for four (not mutually exclusive) tasks, namely

- to document stylized facts that cannot be uncovered by looking at aggregate data for industries or regions,
- to motivate assumptions used in formal theoretical models,
- to test hypotheses derived from theoretical models, and
- to evaluate policy measures.

The following three examples from different areas of economics – firm demography, job creation and destruction, and international firm activities – illustrate the need for, and the research potential of, the use of firm-level data:

- (1) Hopenhayn (1992) considers long run equilibrium in an industry with many price-taking firms producing a homogeneous good. Output is a function of inputs and a random variable that models a firm-specific productivity shock. These shocks are independent across firms and are the reason for the heterogeneity of firms. There are sunk costs to be paid upon entry and entrants do not know their specific shock in advance. Incumbents can choose between exiting or staying in the market. The model leads to three testable hypotheses, namely that firms that exit in year  $t$  were less productive at time  $t-1$  than firms that continue to produce in  $t$ ; that firms that enter in year  $t$  are less productive than incumbent firms in year  $t$ ; and that surviving firms from an entry cohort were more productive than non-surviving firms from this cohort in the start year. Wagner (2007a) uses a panel dataset for all manufacturing plants from

Germany (1995-2002) to test these hypotheses econometrically, and finds that all three hypotheses are supported empirically.

- (2) It is often argued that in Germany jobs are mostly created in small- and medium-sized firms, while large firms generally tend to destroy jobs. The *Mittelstand*, or middle class, is considered the engine of job creation. Using panel data for manufacturing firms, Wagner (2007b) demonstrates that this simple view is wrong. Growing and shrinking firms, entries and exits can all be found in substantial amounts in all size classes within each time period considered. Economic policy measures with a special focus on firms from different size classes, therefore, cannot be justified by pointing to an extraordinary large contribution of these firms to job creation.
- (3) A large number of empirical studies for many countries (surveyed in Wagner 2007c) demonstrate that exporting firms are more productive than non-exporting firms of the same size from the same narrowly defined industry. This stylized fact motivated Melitz (2003) to set aside the standard assumption of homogeneous firms and to develop a model with heterogeneous firms where only the more productive firms in an industry export. This model has become the workhorse of a flourishing body of literature dealing with international firm activities. Using unique, recently released, and nationally representative high-quality longitudinal data at the plant level, Wagner (2007d) presents the first comprehensive evidence on the relationship between exports and productivity for Germany, a leading actor on the world market for manufactured goods. He documents that the positive productivity differential of exporters compared to non-exporters is statistically significant and substantial, even when observed firm characteristics and unobserved firm specific effects are controlled for.

All three examples demonstrate that using firm-level data is not only useful but indispensable for both sound empirical research (including the evaluation of policy measures and the derivation of policy recommendations) and crafting theoretical models that are relevant outside academic journals. In his Nobel lecture, James Heckman (2001, 674) named “the evidence on the pervasiveness of heterogeneity and diversity in economic life” the most important empirical discovery from econometric analyses using microdata. Everybody who ever worked with plant- or enterprise-level data will agree – there is no such thing as a representative firm, not even in 4-digit industries. We would not know this, and would be unable to base our theoretical models and the policy implications derived from these models on this knowledge if firm-level data was not accessible to researchers. Fortunately, such access is possible, as the next section discusses in greater detail.

### 3. Who produces firm-level data, and how can they be accessed by researchers today?

In Germany, the data for establishments and enterprises are collected or constructed by a number of institutions. The most important among them include:

- the Federal Statistical Office (Destatis, *Statistisches Bundesamt*) and the Statistical Offices of the German *Länder* (*Statistische Ämter der Länder*), which administer a large number of surveys as well as secondary statistics;
- the Federal Employment Agency (BA, *Bundesagentur für Arbeit*) and its research institute, the Institute for Employment Research (IAB, *Institut für Arbeitsmarkt- und Berufsforschung*), which uses information on employees covered by social security to construct establishment-level information on the number of employees and their average characteristics, as well as collects information on a wide range of issues for a panel of establishments in annual surveys for the IAB Establishment Panel;
- the Deutsche Bundesbank has a database with information from balance sheets and data on the foreign direct investments of German firms.

Furthermore, firm-level data are collected on a large scale by research institutes (including the Ifo Institute for Economic Research in Munich and the Center for European Economic Research in Mannheim) and by the KfW (*Kreditanstalt für Wiederaufbau*), a bank that is closely related to the German state.

It should be noted that some of these firm-level data include information on the employees working in firms, leading to what is named Linked Employer-Employee (LEE) data. LEE data for Germany are the salary and wage structure surveys (*Gehalts- und Lohnstrukturerhebungen*) from official statistics, and the LIAB, which combines information from the IAB Establishment Panel with employee information from social insurance records.

More information on the firm-level data for Germany, and references to papers describing their information content, are given in Kaiser and Wagner (2008).

In the past, some of the data producers provided access to confidential firm-level data for researchers on the basis of individual contracts and contacts. For example, various statistical offices of the *Länder* allowed researchers to work with firm-level data either via remote data access (i.e., by sending programs to the office, checking their output for violation of data protection rules, and then sending them to the researchers) or by giving them a special status as an unpaid employee, making it feasible for researchers to work with



the microdata inside the office in strictly accordance with all relevant data protection rules. Projects that pursued this type of access formed the network FiDAS<sub>t</sub> – an acronym for Firm-Level Data from Official Statistics (*FirmenDaten aus der Amtlichen Statistik*). Results from these projects are documented in various contributions to professional journals and in three workshop volumes (see Schasse and Wagner 1999; 2001; Pohl et al. 2003). Furthermore, the IAB offered researchers the option of using the data from the IAB Establishment Panel via remote data access and the so-called *Schalterstelle*, a contact person in charge of running the programs and checking the output afterwards (see Kölling 2000).

In recent years, following the suggestions of the German Commission on Improving the Information Infrastructure between Science and Statistics (KVI, *Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik*), most of the important producers of firm-level data – including the Federal Statistical Office and the Statistical Offices of the German *Länder*, the IAB, and the Deutsche Bundesbank – established Research Data Centers that offer researchers convenient ways to work with confidential data via remote data access or by working in-house (see Zühlke et al. 2004; Kohlmann 2005; Lipponer 2003). Furthermore, Scientific Use Files (SUFs) were produced for several datasets that can be used by researchers on their own PCs in the office, as well as Public Use Files (PUFs), which can be used by anybody, including students during courses (see Zwick 2007). Other data producers (like the KfW) offer researchers the opportunity to use the confidential firm-level data in joint projects with employees of the producers, including access to the data while working in-house. A survey of who offers what to whom, and how, is given in Kaiser and Wagner (2008).

Most recently, further progress on the way to a less restrictive access to confidential data was made by locating a Research Data Center outside the data producing institution and inside the institution where the researchers are. The Statistical Office of Berlin and Brandenburg opened a Research Data Center in the building of the German Institute for Economic Research (DIW, *Deutsches Institut für Wirtschaftsforschung*), making the work with the microdata from German official statistics much more convenient for DIW researchers (and for researchers working in the universities nearby).

Compared to twenty, ten, or even five years ago, things have improved a great deal for researchers with regard to access to confidential microdata for establishments and enterprises. As the next section will demonstrate, there is more to come.

#### **4. What will the near future bring? New products in the pipeline**

Compared to firm-level data collected by research institutes, data from official surveys have several advantages: they often cover the whole population of targeted firms (not merely a small sample) and the firms are required to answer and answer correctly (there are no missing cases, no missing values, and – it is to be hoped – no wrong answers). Furthermore, the surveys are usually repeated periodically, and the data from various waves can be combined to build panel datasets. The extra costs associated with preparing data from official surveys for scientific research are not zero, but they are only a tiny fraction of what it would cost to collect data in a new survey. That said, there is one disadvantage of these data from official statistics. Usually, they cover only a small number of items, often fixed by law. This leads to severe limitations with regard to the potential use of these data for scientific analyses.

A promising way to increase the research potential of data from the surveys of official statistics would be to combine the information collected for a unit (enterprise or establishment) in different surveys. This is technically feasible if each unit has a unique identifier (a unit number) that is used in different surveys. Fortunately, this is the case with firms surveyed by the Federal Statistical Office and the Statistical Offices of the German *Länder*. Given that the law allows for matching data from various surveys administered by statistical offices, combined information from these surveys can be used in a single empirical investigation. The following example illustrates how such combinations can increase the research potential of firm-level data from official statistics.

Cost structure surveys collect information on, among other things, turnover and various categories of costs. From these data a rate of return can be computed to proxy the profit situation of the firm. How is this rate of return related to export activities of the firm? This question cannot be answered using these data alone, because no information on exports is collected in the cost structure surveys. Information on exports, however, is available in another survey – a report covering the activities of manufacturing firms, which does not, however, contain any information itself about the profit situation of the firm. Combining data from these two different surveys leads to a dataset that makes it possible to investigate the role of exports for profitability (see Fryges and Wagner 2008).

Matched data from surveys collected by the statistical offices have been used in a number of studies recently. The datasets for these studies have been tailor-made by the Research Data Centers to suit the purposes of each respective study. This is both expensive and time consuming. In the AFiD

project (where AFiD is an acronym for *Amtliche Firmendaten für Deutschland*, or Official Firm-Level Data for Germany) several standardized datasets are prepared that are combinations of data from various surveys (for details see Malchin and Voshage 2009). These combined data are available to researchers via the Research Data Centers of the statistical offices.

Datasets from the AFiD project will offer a convenient way for researchers to investigate questions that could not be answered using data from only one survey. Furthermore, the content of datasets prepared in the AFiD project can be enhanced by adding information from other sources. On the one hand, it is both technically feasible and legal to add data collected in special purpose surveys that are administered by the statistical offices only once. A case in point is the survey on international outsourcing activities of firms recently performed by the German Federal Statistical Office (Statistisches Bundesamt 2008). The data from this survey have a limited amount of information, yet combined with all the other data for firms from the AFiD project, these data offer the opportunity for exciting empirical research on various topics related to the determinants and consequences of international outsourcing. Note that the extra costs of adding these data to the datasets already available are negligible. On the other hand, in accordance with the law, and given that it is technically feasible, data from publicly available sources can be matched with the AFiD data to further enhance the information content of these datasets. To give an example, information about patents granted to the firms can be added. Augmented datasets of this type – or what might be labeled *AFiDplus* data – will offer attractive opportunities for empirical investigations in innovative fields.

While combining information available for a single firm from various surveys done by official statistics (in addition to publicly available information from other sources) in the AFiD project is an attractive way to build new, rich datasets that are worth much more than the sum of their parts to a researcher, even more attractive datasets can be constructed when confidential firm-level microdata from the vaults of different data producers are matched on top of that. To give an example, information on the foreign direct investments of firms is not available from any survey done by the statistical offices, but rather from balance sheet data processed by the Deutsche Bundesbank. Combining AFiD data with the data for foreign direct investments leads to a dataset that makes it possible to investigate problems highly relevant for both scientific analysis and policy debates, including the consequences of foreign direct investments for jobs and wages in Germany.

Due to the sometimes tricky problems related to the definition of economic units, and the different identifiers used for firms by different data producers, this matching can be technically demanding. Furthermore, this is only legally allowed (in Germany in 2008) if each firm explicitly declares in a written statement which of the data it delivers to the different data

producers can be used for the matching. This sets the benchmark fairly high for any project trying to observe this procedure. Recently, the German Federal Ministry of Education and Research (BMBF, *Bundesministerium für Bildung und Forschung*) funded the research project KombiFiD (an acronym for *Kombinierte Firmendaten für Deutschland*, or Combined Firm Data for Germany), a feasibility study in which a large number of firms are asked to agree to match their data and in which the technical problems of matching data across the boundaries of data producers are examined. The data from this feasibility study will be available at the Research Data Centers of the data producers involved in KombiFiD – hopefully beginning in the summer of 2009. More information and up-to-date news on the project can be found at the website: [www.kombifid.de](http://www.kombifid.de).

## 5. A “firm-panelholic’s” wish-list

Even considering all the recent progress that has been made in the way that firm-level data are prepared and made available for the use of independent researchers, and even with all the datasets currently under construction in the projects described above, there are still several wishes left unfulfilled. If a good fairy granted me three wishes related to firm-level data, I would ask for:

- (1) A change in German law allowing the matching of microdata for firms across the boundaries of data producers without requiring written consent from the firms. The reason for this wish is obvious from the discussion presented in section 4.
- (2) Finding ways to combine firm panel data across national borders, and to give researchers access to these data (see the International Public Use Microdata Series Project<sup>2</sup> that collects census data for persons and households from all over the world for a role model dealing with individual level data). The main reason for this wish is the observation that we live in a time of increasing globalization. If the objects of our analysis – the firms – become more and more international, often controlling or being controlled by firms in other countries, the data we use should enable us to learn about the causes and consequences of their behavior by allowing access to micro-level data for all units connected to a firm, legally or otherwise, irrespective of the country these units are located in.

---

2 [www.ipums.org/international](http://www.ipums.org/international).

- (3) Finding ways to enable researchers in Germany to work with firm-level microdata via remote access, available 24 hours a day and 365 days per year, rather than requiring them to send programs to the Research Data Centers or to go there in person (see Hundepool and de Wolf (2005) for a description of a pilot project at Statistics Netherlands). The reason for this wish is obvious to any researcher familiar with the conventional ways of working with confidential firm-level data: while it is possible to work using the current means of access, and it is infinitely better to have this opportunity than not to have any opportunity at all – it remains a second-best solution. Time for research is the ultimate constraint faced by researchers and the means of access available today are extremely time consuming. (As an aside, I would like to add that the SUFs that can be used on the researchers' own PC are in my view no solution when it comes to firm-level data; see Wagner 2005.) While the space limitations for this report make it impossible to go into detail on this point, the example of Denmark (described in Kaiser and Wagner 2008) clearly demonstrates how such an "easy access" policy can be implemented. Based on an approved research proposal, researchers in Denmark can access the data on the mainframe computers in Statistics Denmark from their office PCs, with extremely high penalties for any misuse. Not that long ago, the Kingdom of Denmark began in what is today the northern part of Hamburg, some 40 kilometers north from my office at the Leuphana. Given the high price of beer in Denmark I am not sure that I would wish this still to be the case – yet when I look at the ease of access to all kinds of confidential microdata that my colleagues at Danish universities enjoy, I do feel some regret. So, at the end of the day, I do wish that we would start to learn from the Danish experience.

## References:

- Fryges, H. and Wagner, J. (2008): Exports and profitability – First evidence for German manufacturing firms. [Unpublished manuscript].
- Heckman, J.J. (2001): Micro Data, heterogeneity, and the evaluation of public policy: Nobel lecture. *Journal of Political Economy* 109 (4), 673-748.
- Hopenhayn, H. (1992): Entry, exit, and firm dynamics in long run equilibrium. *Econometrica* 60 (5), 1127-1150.
- Hundepool, A. and de Wolf, P.-P. (2005): OnSite@Home: Remote Access at Statistics Netherlands. [Unpublished manuscript].
- Kaiser, U. and Wagner, J. (2008): Neue Möglichkeiten zur Nutzung vertraulicher amtlicher Personen- und Firmendaten. *Perspektiven der Wirtschaftspolitik* 9 (3), 329-349.
- Kölling, A. (2000): The IAB-Establishment Panel. *Schmollers Jahrbuch* 120 (2), 291-300.
- Kohlmann, A. (2005): The Research Data Center of the Federal Employment Service in the Institute for Employment Research. *Schmollers Jahrbuch* 125 (3), 437-447.
- Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik (KVI) (Ed.) (2001): Wege zu einer besseren informationellen Infrastruktur. Baden-Baden.
- Liebig, S. (2009): Organizational Data. RatSWD Working Paper No. 67.
- Lippuner, A. (2003): Deutsche Bundesbank's FDI micro database. *Schmollers Jahrbuch* 123 (4), 593-600.
- Malchin, A. and Voshage, R. (2009): Official Firm Data for Germany. *Schmollers Jahrbuch* 129 (3), 501-513.
- Melitz, M.J. (2003): The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica* 71 (6), 1695-1725.
- Pohl, R./Fischer, J./Rockmann, U. and Semlinger, K. (Eds.) (2003): Analysen zur regionalen Industrieentwicklung. Sonderauswertungen einzelbetrieblicher Daten der Amtlichen Statistik. Berlin.
- Schasse, U. and Wagner, J. (Eds.) (1999): Entwicklung von Arbeitsplätzen, Exporten und Produktivität im interregionalen Vergleich – Empirische Untersuchungen mit Betriebspaneldaten. Hannover.
- Schasse, U. and Wagner, J. (Eds.) (2001): Regionale Wirtschaftsanalysen mit Betriebspaneldaten – Ansätze und Ergebnisse. Hannover.
- Statistisches Bundesamt (2008): Verlagerung wirtschaftlicher Aktivitäten. Erste Ergebnisse. Wiesbaden.
- Wagner, J. (2005): Anonymized firm data under test: Evidence from a replication study. *Jahrbücher für Nationalökonomie und Statistik* 225 (5), 584-591.
- Wagner, J. (2007a): Entry, exit and productivity. Empirical results for German manufacturing industries. University of Lüneburg Working Paper Series in Economics 44, March.
- Wagner, J. (2007b): Jobmotor Mittelstand? Arbeitsplatzdynamik und Betriebsgröße in der westdeutschen Industrie. *Vierteljahrshefte zur Wirtschaftsforschung* 76 (3), 76-87.
- Wagner, J. (2007c): Exports and productivity: A survey of the evidence from firm-level data. *The World Economy* 30 (1), 60-82.

- Wagner, J. (2007d): Exports and productivity in Germany. *Applied Economics Quarterly* 53 (4), 353-373.
- Zühlke, S./Zwick, M./Scharnhorst, S. and Wende, Th. (2004): The research data centers of the Federal Statistical Office and the statistical offices of the Länder. *Schmollers Jahrbuch* 124 (4), 567-578.
- Zwick, M. (2007): CAMPUS files – Free public use files for teaching purposes. *Schmollers Jahrbuch* 127 (4), 655-668.





## 5. STATE, FAMILY, AND HEALTH

### 5.1 Public Finance

*Thiess Büttner\**

---

\* The author is grateful to U. Rehfeld for helpful comments on an earlier draft. All errors remain the author's responsibility. The support of the German Data Forum (RatSWD) is gratefully acknowledged.

Contact:

Thiess Büttner  
Ifo Institute for Economic Research  
Poschingerstr. 5  
81679 Munich  
Germany  
e-mail: [buettner\[at\]ifo.de](mailto:buettner[at]ifo.de)

## **Abstract**

This paper briefly surveys the available data sources relevant to the empirical study of public finance in Germany and discusses future developments. It starts from the notion that public finance deals with decisions made by diverse agents, not only by different levels of government but also private households and firms. As a result empirical research requires different types of data. Budgetary statistics capture government decisions to some extent, although these statistics have shortcomings related to the quality of public service provisions and the revenue instruments. In order to study the decisions made by the other agents, individual-level data is also required. While there has been some recent progress in this direction, the combination of various datasets at the individual level is a key priority.

Keywords: empirical research, public finance, budgetary statistics, revenue statistics, micro-level data, taxpayer data

## **1. Research questions**

Public finance is an area concerned with decisions made by collective agents and institutions with the impact of those decisions bearing on the economy and individual agents. This definition implies that, depending on the specific topic being researched, disparate types of data are used and may have to be combined.

Empirical research on public finance traditionally addresses the decision making of the public sector itself. This includes the substantial efforts that have been undertaken to monitor developments in the public sector in terms of the budget as well as with regard to service provision and inputs. In most countries the public sector shows a marked vertical and horizontal structure such that research addresses all levels of government including national and federal governments, state governments, and local governments, as well as separate bodies such as school districts and public enterprises.

With regard to policy areas, research on public finance tends to take a comprehensive view. In areas where policy implementation takes specific forms and includes private institutions, the research has developed into autonomous subfields. This includes areas such as Health and Education. While some research in these areas has close connections to public finance in the general sense, this paper will not discuss these specialized areas of empirical research infrastructure.

A significant part of empirical research in public finance focuses on the impact of instruments of public policy on the economy – including taxes and various types of government spending such as subsidies and transfers. A particular focus is on the impact of those instruments on individual agents,

such as households and firms. This usually requires data about individual agents. At the same time, however, the study of long-term implications of policies requires not only cross-sectional, but also longitudinal data.

Due to its intrinsic complexity, research on taxation is often concerned with simulating rather than testing for tax effects. The corresponding simulation models need detailed information about the various components of the tax accounts and, hence, are ideally based on micro-level data for taxpayers.

Given the large share of resources expended by the public sector, empirical research is also concerned with the macroeconomic consequences of fiscal policies at an aggregate level such as the national, regional, or state levels. There is also empirical research that aims to provide a comprehensive picture of the economic consequences of government policies using general-equilibrium simulation models. These models require detailed information not only about the public sector but also about household and firm sectors and may also utilize input-output tables. Moreover, these models usually employ various parameters that originate in previous empirical research.

Empirical research on public finance, however, must not only concern itself with different types of agents, it also faces specific measurement problems with regard to government decisions and policies. Often, research is concerned with data on government spending or revenues. However, in many circumstances empirical research benefits from using more detailed information on specific government policies. This is particularly applicable in the context of taxation where the relevant policies are concerned with determining very specific parameters such as statutory tax rates, tax brackets, or tax incentives. Generally, a detailed knowledge of the law and its implementation is required. Measurement issues are also important with regard to the supply side of the public sector where the use of expenditure data is often not sufficient to capture government policies if the analysis is concerned with public service provision. As quantity and quality of public service provision are difficult to capture, research often must resort to using survey data where respondents assess the supply of public services.

## **2. Status quo: Databases and access**

Given the various types of research questions, a useful way to structure a discussion of databases is to distinguish the agents whose decisions are under consideration as well as the type of policy under consideration such as taxes, public service provision, or social policies.

## 2.1 Data on governments

The basic data source for empirical research are the fiscal accounts that capture expenditures and revenues, as well as information on stocks such as government debt or assets. With regard to Germany, the Federal Statistical Office and the Statistical Offices of the German *Länder* offer a broad set of detailed data covering public spending, detailed by types of expenditure and categories of revenue. These statistics refer to the various fiscal tiers (federal, state, county, and municipality) and some important *parafisci* (para-fiscal organizations), such as social insurance. Quarterly data are provided about three months after the end of the respective quarter. Detailed data on government functions are based on the annual accounts available about two years after the respective year. Budget information is augmented by aggregate tax revenue statistics that report revenue pertaining to specific taxes. The standard set of statistics also includes information about the stocks of debt. The monitoring of government activities by the statistical offices further includes information about employment in the public sector. Finally, information about the annual accounts of state-owned enterprises is also available.

Data access is easy for federal- and state-level data as well as for the consolidated budget of the public sector: in all cases information is available on the website of the Federal Statistical Office. However, with regard to the state level, not all statistics are provided as part of the standard program. For instance, detailed data regarding both the type of expenditure and the function of government require separate requests. While a virtue of the German system of fiscal federalism is that the fiscal classification used for counties and municipalities is almost identical across the German states, or *Länder*, detailed data on government below the state level are only available for individual states subject to the approval of the Statistical Office in that respective state.<sup>1</sup>

Given the large share of public expenditure allocated to social welfare, several statistics of the statistical offices focus on specific programs, and the Federal Ministry of Labor and Social Affairs (BMAS, *Bundesministerium für Arbeit und Soziales*) provides even more comprehensive statistics (*Sozialbudget*).

Government statistical offices hold individual tax files for several major taxes that provide more detailed information and are made available for research as Scientific Use Files. Combination into panel data is possible if based on tax identifiers. However, with most other taxes the data are triennial starting with 1992 or 1995. Data access is restricted to the Research Data Centers at the national and state level and may be further restricted by remote processing of routines.

---

1 An exception is a study by Borck et al. (2007), which uses a comprehensive dataset for all German municipalities.

While the fiscal variables refer to the executive branch of the government, information about the legislature is also provided by the statistical offices. At the website <http://www.bundeswahlleiter.de>, detailed data on the results of federal, state, and local elections can be downloaded. Information about annual and medium-term budget planning is provided through the Federal Ministry of Finance (BMF, *Bundesministerium der Finanzen*) or through the Finance Ministries of the German *Länder* (*Finanzministerien der Länder*). Data on auditing is available for the federal and the state level only. Results from the auditing of lower-level governments are generally not available.

With regard to quantity and quality of public services provided by the various governments, the data supplied by the statistical offices are rather limited. Specific statistics exist for some functions of government. For instance, statistics on higher education provide data on enrollment at universities and the universities of applied sciences (*Fachhochschulen*) according to university and field of study and include additional information about the background of students. However student test scores as assembled in the OECD's PISA initiative are not provided at the state level in a way that allows meaningful cross-state comparisons. Since education is the key responsibility of the state governments in Germany, this restriction is a severe limitation for empirical research in Germany.

At the municipal level, the *Statistical Yearbook of German Communes* (edited by the German Association of Cities) provides some further information on the supply of government services, however, this data focuses on larger cities. Research that is concerned with the supply of public services at the local level might need to resort to survey data, where, however, the number of respondents is often small. An exception is the "Perspektive Deutschland" where waves four (2004/2005) and five (2005/2006) provide survey responses for several aspects of local living conditions including public services at county level. Data access is provided through the GESIS<sup>2</sup> Data Archive in Cologne.

## 2.2 Household data

An important part of empirical research in public finance is concerned with the impact of public policies on household decisions such as consumption, labor supply, or location. For this purpose, all sorts of household data are used such as the German Socio-Economic Panel (SOEP, *Sozio-oekonomisches Panel*), Mikrocensus, or the Income and Consumption Survey (EVS, *Einkommens- und Verbrauchsstichprobe*). The latter is particularly interes-

---

2    Leibniz Institute for the Social Sciences (GESIS, *Leibniz-Institut für Sozialwissenschaften*).

ting as it offers some direct information about taxes paid whereas the SOEP employs imputed values (Becker et al. 2002).

Given the importance of the specific institutional details of the tax code, research often uses taxpayer panels (see below) even if these have limited information about household characteristics. This includes the *IAW-Einkommensteuerpanel* (Gottfried and Schellhorn 2001) that builds on individual tax information in the state of Baden-Wuerttemberg. More recently, the Federal Statistical Office has begun to provide an annual taxpayer panel, a project that makes taxpayer data available on an annual basis beginning in 2001 (see Kriete-Dodds and Vorgrimler 2007). While the taxpayer panel can only be used via controlled remote processing, the triennial micro-level income tax statistics is available as a Scientific Use File (FAST) in the Research Data Center of the Statistical Offices of the German *Länder*.

Another important area of research is concerned with the consequences of social policy on individual choice. However, the key issues in this context such as distribution or labor market participation suggest that the discussion of data availability and the conditions for empirical analysis are best addressed in the context of poverty (see Hauser, 2009) and labor market research (see Bender and Möller 2010; Schneider 2010).

### 2.3 Firm-level data

To study the impact of government policies on firm decisions, a large body of research utilizes firm-level data that capture investment, financial structure, and many other dimensions of firm decisions. However, financial statement data as provided by Hoppenstedt or Creditreform (DAFNE) usually report tax payments that capture not only the tax burden or tax incentives but also reflect firm performance and/or tax planning. The resulting problem of the endogeneity of tax variables has made it difficult to identify the role of the tax system for investment or the financial structure of firms. Rather than using tax payments, research might exploit differences between firms that lead to differences in taxation due to the specifics of the tax law, perhaps related to legal form or firm size.

Over the last decade, however, empirical research has been more successful in addressing these issues by employing data for multinational firms operating in different countries. The advantage here is that policies including tax policies show marked variation across countries that can be exploited for identification purposes. As a consequence, a great deal of research has been concerned with multinational data. Financial statement data for German and European multinationals or multinationals operating in Europe are provided by commercial providers such as Bureau von Dijk's Amadeus database. A unique data source for studying multinationals is the Bundesbank's MiDi database that currently provides annual firm-level panel

data for the period 1996 to 2004. The collection of the data is prescribed by German law, which determines reporting mandates for international transactions (Lipponer 2006). A shortcoming of the MiDi dataset is that it provides limited information about the parent companies of domestic affiliates and German parents of foreign affiliates.

Alternatively, research in this area is concerned with data that exploits institutional variations across regions. In the context of company taxation, for instance, many studies exploit the local variation in the local business tax (*Gewerbesteuer*). Research opportunities are provided by the corporate balance sheet database (*Jahresabschlussdaten*) of the Deutsche Bundesbank (Stoess 2001), available within the bank's research center, by the IAB establishment panel (*IAB-Betriebspanel*), where Scientific Use Files are provided through the Research Data Center of the Federal Employment Agency at the Institute for Employment Research (IAB, *Institut für Arbeitsmarkt- und Berufsforschung* within the BA, *Bundesagentur für Arbeit*), and by the taxpayer data for the local business tax. The latter is a triennial micro-level dataset that currently provides information for 1998 and 2001. It can be accessed within the Research Data Centers of the Statistical Offices of the German *Länder*, though data access is further restricted by controlled remote processing.

Given the difficulties in identifying tax effects, research is also concerned with setting up simulation models. As with the analysis of household decisions, micro-level tax statistics are particularly helpful for this purpose. In addition to the local business tax, the triennial micro-level tax statistics for the corporation tax, and, with regard to the unincorporated firms, also the personal income tax statistic can be used – all of them are provided within the Research Data Center of the Statistical Offices.

## 2.4 Tax policy & institutional data

Typical of a large part of empirical research in public finance is a detailed characterization of institutions, including tax systems, social security systems, or specific laws that govern government policies. At the international level, however, there are some supra-national bodies such as the OECD and the European Commission that provide data on tax systems and institutional characteristics of countries such as the vertical structure of the public sector.

Generally speaking, data collection is easier for subnational entities. For instance, information about the tax burden associated with the local business tax (*Gewerbesteuer*) and the land tax (*Grundsteuer*) is provided by the



statistical offices, at least at the level of counties.<sup>3</sup> However, the effective tax burden on land is not known due to the substantial discrepancy between market value of land and the assessed value. While land has been assessed for the estate and gift tax (*Erbschaftsteuer*) according to market prices since the mid-1990s, information about the assessment is not provided by the statistical offices.

### 3. Future developments

An important issue for the future development of research infrastructure is the combination of diverse data sources. This refers in particular to corporate and personal income taxation where taxpayer panel data so far have not been merged across different taxes. However, even at the level of the corporation, the tax burden consists of local business taxes as well as of corporate taxes. Moreover, a combination of taxpayer data with other firm-level data could yield substantially improved datasets where firm decisions as well as firm-specific conditions could be modeled much more precisely. The recent KombiFiD initiative of the Federal Statistical Office, the IAB, and the Deutsche Bundesbank, which aims at providing those combined datasets is to be greatly appreciated (see Bender et al. 2007). In particular, the combination with the Bundesbank MiDi data would vastly improve conditions for empirical research. Another promising data combination project is the Economics & Business Data Center (EBCD) initiative of Munich University in collaboration with the ifo Institute, which aims to combine ifo-firm survey data (Becker and Wohlrabe 2008) with commercial financial statement data including Amadeus and Hoppenstedt. An interesting aspect of this project is that it relies on randomized record linkage.

---

3 The common practice is to report weighted averages of municipal tax rates where the weights correspond to the municipal government's tax revenue. However, this practice is problematic. To see this, consider the weighted collection rate (*Hebesatz*) for a set of municipalities:

$$\frac{\sum_{i=1}^N h_i G_i}{\sum_i G_i}$$

where  $G_i$  is the revenue at a standardized collection rate (*Grundbetrag*). Since  $G_i$  is a declining function of the own collection rate and an increasing function of the tax rates of other municipalities, municipalities with high tax rates tend to receive a smaller weight. As a consequence, using the weighted average tends to yield biased results: tax increases are underestimated, tax decreases are overestimated.

A data combination that would help to address important issues pertaining to the tax system would be the combination of individual and firm-level taxpayer data. This would be an important step towards creating a reliable and comprehensive empirical basis for research on tax policy and reforms. The BMF has started an initiative in this direction. In this context, it should be noted that the growing complexity of tax issues has led other countries to set up micro-simulation models that are used for revenue estimation and also revenue forecasting purposes. However, currently no such attempts have materialized in Germany.

With regard to the analysis of the finances of subnational jurisdictions, the traditional financial accounting system is subject to change. As of 2009, North Rhine-Westphalia, the largest German state, introduced a new system of accounts, entitled “Neues Kommunales Finanzmanagement” (NKF), which replaced the current cash-based accounting. While this might systematically improve the information about controlled enterprises and liabilities, some new problems regarding assessment and valuation are coming up. Moreover, the data series will suffer from an important structural break.

As the current analysis of municipal finances has to resort to the last 1987 census, another important development for empirical research at the local level is the new census planned for 2011. While this update is important, it should be noted that current access to census data at the municipal level is difficult, since data access is restricted by the Statistical Offices of the *Länder*.

#### **4. Future developments: European and international challenges**

A large amount of research in public finance is concerned with the consequences of international economic and political integration for public policies. This includes cross-border flows not only of goods and services but also migration, factor movements, capital flows, and the emergence of multinational enterprises. Even if those cross-border issues are of particular importance in the European context, there is only very limited information available. Research so far has centered around specific datasets, many of which are subject to important qualifications. To merge those datasets with other more standard datasets of households and firms would substantially improve the conditions for empirical research using German data. Therefore, initiatives like KombiFiD are very welcome.

## 5. Conclusions and recommendations

Empirical research in public finance that aims at monitoring and assessing budgetary performance has access to a rich body of financial accounts that enables researchers to assess federal and state budgets. While the statistics for lower-level governments are available, data access is unduly restrictive. Since there is no justification to hold back financial data on these governments, the statistical offices should rethink their publication strategy.

With regard to statistics on major taxes, the Statistical Offices of the German *Länder* and the Federal Statistical Office have recently improved conditions for empirical research by providing micro-level taxpayer data. This is a major achievement. However, there are serious restrictions in data access and also the limited information about the background of tax payers in these statistics constitute a significant obstacle for the exploitation of this data.

Attempts to combine different data for research purposes are greatly appreciated. However, data combination should not be confined to providing firm-level data; it is also important to combine different taxpayer statistics in order to get a comprehensive and consistent data source for the empirical analysis of the tax system.

With regard to research on the supply side of public services, there is inadequate data availability. To some extent this reflects the fundamental problem of measuring public services – expenditure data offers only very limited information about the quantity and quality of public service provision. There has also been some progress made in specific areas of government policies such as social policy. However, in other areas, such as public education, the information about quality that is available in principle has not been made available for research.

## References:

- Becker, I./Frick, J./Grabka, M.M./Hauser, R./Krause, P. and Wagner, G.G. (2002): A Comparison of the Main Household Income Surveys for Germany: EVS and SOEP. In: Hauser, R. and Becker, I. (Eds.): Reporting on Income Distribution and Poverty. Perspectives from a German and European Point of View. Heidelberg.
- Becker, S.O. and Wohlrabe, K. (2008): Micro Data at the Ifo Institute for Economic Research – The “Ifo Business Survey” Usage and Access. *Schmollers Jahrbuch* 128 (2), 307-319.
- Bellmann, L. (2002): Das IAB-Betriebspanel: Konzeption und Anwendungsbereiche. *Allgemeines statistisches Archiv* 86 (2), 177-188.
- Bender, S. and Möller, J. (2010): Data from the Federal Employment Agency. [In this publication].
- Bender, S./Wagner, J. and Zwick, M. (2007): KombiFiD – Kombinierte Firmendaten für Deutschland. FDZ-Arbeitspapier Nr. 21.
- Borck, R./Caliendo, M. and Steiner, V. (2007): Fiscal Competition and the Composition of Public Spending: Theory and Evidence. *FinanzArchiv* 63 (2), 264-277.
- Gottfried, P. and Schellhorn, H. (2001): Das IAW-Einkommensteuerpanel und das Mikrosimulationsmodell SIMST. IAW-Diskussionspapier. Tübingen.
- Hauser, R. (2009): Household Income, Poverty and Wealth. RatSWD Working Paper No. 53.
- Kriete-Dodds, S. and Vorgrimler, D. (2007): Das Taxpayer-Panel der jährlichen Einkommensteuerstatistik. *Wirtschaft und Statistik* 2007 (1), 77-85.
- Lipponer, A. (2006): Microdatabase direct investment – MiDi. A brief guide. Bundesbank Working Paper.
- Schneider, H. (2010): More and Better Data for Labor Market Research. Proposals for Efficient Access to the Currently Unused Potential of Official Statistical Data. [In this publication].
- Stoess, E. (2001): Deutsche Bundesbank's Corporate Balance Sheet Statistics And Areas of Application. *Schmollers Jahrbuch* 121 (1), 131-137.

## 5.2 Household Income, Poverty, and Wealth

*Richard Hauser\**

---

\* This advisory report reflects the state of affairs as at 2009.

Contact:

Richard Hauser  
Goethe University  
Department of Economics and Business Administration  
POB 11 19 32  
60054 Frankfurt am Main  
Germany  
e-mail: [r.hauser\[at\]em.uni-frankfurt.de](mailto:r.hauser[at]em.uni-frankfurt.de)

## Abstract

This paper concentrates on the official statistics on household income, poverty, and wealth. It characterizes the main research questions in this field, and presents an overview of the available statistics and Scientific Use Files produced by the four Research Data Centers in Germany: the Research Data Center of the Federal Statistical Office, the Research Data Center of the Statistical Offices of the German *Länder*, the Research Data Center of the Federal Employment Agency at the Institute for Employment Research (IAB, *Institut für Arbeitsmarkt- und Berufsforschung* within the BA, *Bundesagentur für Arbeit*), and the Research Data Center of the German Pension Insurance (RV, *Deutsche Rentenversicherung*). We support the recommendations of a peer review group for the Federal Statistical Office based on the European Statistics Code of Practice, and suggest peer reviews for all data-producing bodies including ministries. We repeat a recommendation of a former commission to find ways of distributing Scientific Use Files to reliable foreign research institutes. Special recommendations refer to the improvement of survey methods and extended questionnaires of the Income and Consumption Survey (EVS, *Einkommens- und Verbrauchsstichprobe*) and the German contribution to the European Statistics on Income and Living Conditions (EU-SILC). We also recommend a harmonization of the administrative statistics on the various minimum benefit programs, and the development of a single Scientific Use File for all minimum benefit recipients.

Keywords: Research Data Center, Scientific Use Files, household income, wealth, minimum benefits, EVS, EU-SILC

## 1. Introduction

Individual and household well-being is strongly determined by income and wealth. The levels and distributions of these economic resources and their changes within a society over time are of utmost importance from a scientific and political point of view. A minimum amount of regular income is also a necessary although not always a sufficient condition for avoiding poverty. Strictly speaking, net equivalent income and net wealth are the main variables of interest here.<sup>1</sup> To calculate these variables, however, one needs information on all sources of income and all components of wealth, as well as on all personal taxes and social security contributions.<sup>2</sup>

---

1 Net equivalent income is a weighted per capita income derived from the net income of the individual's household. Net wealth is the difference between all of a household's assets and its debts.

2 The most important research questions and the current status of research can be gathered from Atkinson, A.B. and Bourguignon, F. (Eds.) (2000): *Handbook of Income Distribution*, vol 1. Amsterdam et al. The historical perspective on changes in the share of high incomes is presented in Atkinson, A.B. and Piketty, T. (Eds.) (2007): *Top Incomes over the 20th*

International comparisons of income and wealth distributions and of the size and composition of the population in poverty are in increasingly high demand both in the European Union<sup>3</sup> and worldwide.<sup>4</sup> These require comparable definitions of the variables measured in national surveys and administrative datasets.<sup>5</sup>

Income and wealth distributions are “anonymous” in the sense that the overall distributions do not change if individuals or households simply switch places in the distribution. When looking at the distributions of income and wealth from an individual point of view, however, one is also able to identify changes in the relative positions of concrete individuals in the income and wealth hierarchy. This is especially important for the analysis of changes in the composition of the poor population due to ascents out of and descents into income poverty.<sup>6</sup> If one extends the perspective from a short-run to a long-run view, it becomes possible to identify changes in individuals’ economic resources over the entire life course. Part of an individual’s life course may be as a member of a family, meaning that the life courses of other family members and their interrelationships are also of interest. Moreover, information about private transfers between households and intergenerational transfers of income and wealth (gifts, inheritances) is needed to gain a complete picture.

The first step of an analysis in the field of income, wealth, and income poverty is always to *describe* the present situation based on household income and wealth statistics. An even greater challenge, however, is to analyze the factors that have produced the existing distribution and that will cause

---

Century, A Contrast Between continental European and English-Speaking Countries, Oxford. An overview of wealth distributions is provided by Wolff, E.N. (Ed.) (2006): International Perspectives on Household Wealth, Cheltenham/UK and Northampton/MA, US. Problems of poverty are summarized in Huster, E.-U./Boeckh, J. and Mogge-Grotjahn, H. (Eds.) (2008): Handbuch Armut und Soziale Ausgrenzung. Wiesbaden; and in Jenkins, S.P. and Micklewright, J. (Eds.) (2007): Inequality and Poverty Re-examined. Oxford.

3 The EU has defined the so-called Laeken indicators which have to be calculated regularly by each member state to facilitate comparisons between its members.

4 Compare OECD (2008): Growing Unequal? Income Distribution and Poverty in OECD Countries, Paris.

5 See Expert Group on Household Income Statistics (The Canberra Group) (2001): Final Report and Recommendations, Ottawa. The Luxembourg Income Study (LIS) collects data on income and wealth for about thirty countries and takes great effort to make them comparable based on these recommendations.

6 A panel study on social assistance recipients gave rise to new insights. See Leisering, L. and Leibfried, S. (1999): Time and Poverty in Western Welfare States, United Germany in Perspective. Cambridge. Many studies on income mobility use data from the German Socio-Economic Panel (SOEP, *Sozio-oekonomisches Panel*), a social science based panel that is located at the German Institute of Economic Research (DIW Berlin, *Deutsches Institut für Wirtschaftsforschung*). Comparative studies of income mobility can be carried out based on the Comparative National Equivalent File (CNEF) that presently comprises the national panels of five countries (Germany, United Kingdom, Canada, Australia, The Netherlands). It is organized and distributed by Cornell University, Ithaca, NY, US.



changes in this distribution and in the relative positions of individuals in the income and wealth hierarchy, especially those in poverty. Our objective, therefore, is to find *explanations* and to make *predictions*. Although there is no comprehensive theory of the personal distribution of income and wealth, one can say that it results from an interaction among macroeconomic and demographic trends, institutional arrangements, and personal characteristics. Social and fiscal policy decisions that change the institutional arrangements work within this general setting. While information on macroeconomic and demographic developments, institutional arrangements, and policy decisions has to be obtained from other sources, information on relevant personal characteristics should be contained in the same data file with information on individual or household income. This is necessary in cross-sectional household surveys as well as in household panel surveys. While simulations of the first-round effects of social and fiscal policy changes usually neglect behavioral responses by assumption, the prediction of second- and third-round effects requires estimates of individual behavioral responses with respect to working time, consumption and savings, and changes in the portfolio structure of wealth holdings. Econometric estimates of these behavioral responses should therefore be based on variables contained in the same dataset as the income and wealth variables. Usually, however, one has to ignore the macro-level consequences of micro-level behavioral changes due to the lack of an integrated micro-macro model.

Given the current state of research, in the following we will examine the sources of official statistical data currently available for analyses of income, wealth, and poverty. On this basis, we will formulate recommendations for improving specific components of the information infrastructure in Germany. Cross-sectional and longitudinal household surveys conducted independently by social science research organizations will not be dealt with.

## **2. An overview of public data sources and their availability for research on income, wealth, and poverty**

Since 2001, when the German Commission on Improving the Information Infrastructure between Science and Statistics (KVI, *Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik*) published its first set of recommendations, the German statistical infrastructure for empirical research in the economic and social sciences has improved dramatically.<sup>7</sup> For research on the distribution of income and

---

<sup>7</sup> Towards an Improved Statistical Infrastructure – Summary Report of the Commission set up by the Federal Ministry of Education and Research to Improve the Statistical

wealth and of income poverty, the following Scientific Use Files are provided by public institutions through various Research Data Centers.<sup>8,9</sup>

## 2.1 Research Data Centers of the Federal Statistical Office and the Statistical Offices of the German Länder<sup>10</sup>

These Research Data Centers offer a range of Scientific Use Files and provide researchers in many fields diverse possibilities for working either on-site or via remote computing. In the following, only those files are listed that refer to income, wealth, and poverty.

- (1) Microcensus (1973-2006)
- (2) Income and Consumption Surveys (EVS, *Einkommens- und Verbrauchsstichprobe*) (1962/63, 1969, 1973, 1978, 1983, 1988, 1993, 1998, 2003)
- (3) German contribution “Leben in Deutschland” to the European Statistics on Income and Living Conditions (EU-SILC) (2005), European Community Household Panel (ECHP) (1994, 1995, 1996)
- (4) Income Tax Files (1992, 1995, 1998, 2001, 2004)
- (5) Taxpayer Panel (2001, 2002, 2003, 2004)
- (6) Inheritance Tax File (2002)
- (7) Social Assistance Files (1998-2004)

Datasets (1), (2), and (3) are sample surveys, while (4), (5), (6), and (7) are samples of administrative datasets. All these datasets are relevant for analyses of income, wealth, and poverty, but we will only comment on the surveys. The relatively new tax files are very promising for distributional analyses.<sup>11</sup> Methodological research on problems in administrative datasets is still in progress.

---

Infrastructure in Cooperation with the Scientific Community and Official Statistics (KVI), reprinted in: Schmollers Jahrbuch 121 (3), 443-468.

8 Details of the four Research Data Centers are described in several articles in the volume Rolf, G./Zwick, M. and Wagner, G.G. (Eds.) (2008): Fortschritte der informationellen Infrastruktur in Deutschland. Festschrift für Johann Hahlen zum 65. Geburtstag und Hans-Jürgen Krupp zum 75. Geburtstag. Baden-Baden.

9 In addition to the distribution of Scientific Use Files, the Research Data Centers also provide workplaces for guest researchers on site and facilities for remote computing with all the surveys mentioned.

10 [www.Forschungsdatenzentrum.de](http://www.Forschungsdatenzentrum.de).

11 See Bach, St./Corneo, G. and Steiner, V. (2007): From Bottom to Top: The Entire Distribution of Market Income in Germany, 1991-2001. DIW Discussion Paper No. 683. And Bach, St./Corneo, G. and Steiner, V. (2008): Effective Taxation of Top Incomes in Germany, 1992-2002. DIW Discussion Paper No. 767.

2.2 *Research Data Center of the Federal Employment Agency at the Institute for Employment Research (IAB, Institut für Arbeitsmarkt- und Berufsforschung within the BA, Bundesagentur für Arbeit)*

- (8) IAB Employment Sample (*IAB-Beschäftigtenstichprobe*) (1975-2004)
- (9) Cross-Sectional Survey “Life Situation and Social Security” (LSS 2005, *Querschnittsbefragung Lebenssituation und Soziale Sicherung*)
- (10) Panel Study “Labor Market and Social Security” (PASS, *Panel Arbeitsmarkt und soziale Sicherung*);
- (11) BA-Employment Panel (BAP, *BA Beschäftigtenpanel*);
- (12) Integrated Employment Biographies Sample (IEBS).

Dataset (8) is a valuable administrative dataset for research on the distribution of individual labor income, which can contribute to explaining the distribution of net equivalent income derived from household net income. These statistics only contain incomes of higher-earning employed individuals up to the ceiling on social security contributions. Dataset (9) is an individual survey that can be used for analyses of income distribution among households, but its comparability with other surveys is limited. Datasets (10) and (11) are based on surveys focusing on the long-term unemployed and thus can contribute to a partial explanation of net household income and of poverty, but only for this subgroup of the population. Dataset (12) is a sample from a longitudinal administrative data survey.

2.3 *Research Data Center of the German Pension Insurance (RV, Deutsche Rentenversicherung)*

- (13) Cross-sectional files for the years 2003, 2004, 2005, and 2006 of the German Pension Insurance on pensions in payment (*Rentenbestand*), on new pensions awarded (*Rentenzugang*), on cessation of pension payment (*Rentenwegfall*), and on actively (currently) insured persons (only 2004 and 2005)
- (14) Cross-section files on special topics and groups:
  - a) persons with reduction/loss of earnings capacity and their diagnosis (2003, 2004, 2005, 2006)
  - b) Scientific Use File with reduced information on pensions in payment, newly awarded pensions, cessation of pensions (1993-2005)
  - c) Scientific Use File on the qualifications of persons with newly awarded pensions (2003).

These various Scientific Use Files can only be used to describe the distribution of pensions by case and to explain their levels as determined by the pension formulae. Since individuals may receive more than one pension – from the German Pension Insurance as well as from other old age protection systems, these datasets are not sufficient to estimate the total pension income of elderly individuals. For this, one needs household surveys that record all types of old age income. Although there exist several very good household surveys on income of the elderly and even on pension entitlements accrued for persons over 40 (ASID, *Alterssicherung in Deutschland* 1986, 1995, 2003, and AVID, *Altersvorsorge in Deutschland* 1996, 2005), they are not available for independent scientific research but only for research commissioned by the Federal Ministry of Labor and Social Affairs (BMAS, *Bundesministerium für Arbeit und Soziales*).<sup>12</sup> This contradicts the recommendations of the 2001 KVI report that all surveys paid for by public money should be available to researchers.

### **3. Problems with the information infrastructure for research in income, wealth, and poverty provided by public bodies**

#### *3.1 Problems mentioned in a report by a European peer review group*

The peer review initiated by Eurostat identified problems with the existing statistics of the German Statistical Offices and made recommendations for improvement.<sup>13</sup> This peer review, based on the *European Statistics Code of Practice*, dealt with how the German Statistical Offices produce their official statistics, internal organization, quality control, and the distribution of statistics to the public and the research community.<sup>14</sup> Some of the problems detected also apply to the quality of data distributed for scientific research.<sup>15</sup> The recommendations for improvement contained in the report implicitly

12 Results of AVID 2005 are published by: Frommert, D./Ohsmann, S. and Rehfeld, U.G. (2008): *Altersvorsorge in Deutschland 2005 (AVID 2005) – Die neue Studie im Überblick*. Deutsche Rentenversicherung 63 (1), 1-19. A critique of these results can be found in: Hauser, R. (2007): *Altersarmut unterschätzt*. Soziale Sicherheit, Zeitschrift für Arbeit und Soziales 56 (12), 416-419.

13 Peer review on the implementation of the European Statistics Code of Practice, Country visited: Germany, by Geert Bruinooge (statistics Netherlands, Daniel Defays (Eurostat), Paloma Seoane Spiegelberg (INE, Spain), March 10, 2008 (available at [www.destatis.de](http://www.destatis.de)).

14 The Research Data Centers of the Federal Statistical Office and of the Statistical Offices of the German *Länder* were accepted as “best practice.”

15 See Peer review, section 7: principle 4 (quality commitment), principle 7 (sound methodology), principle 11 (relevance), principle 12 (accuracy and reliability), and principle 15 (accessibility and clarity).

indicate where the problems lie. The following are key elements of five of these recommendations:

- Depending on resources available and taking into account the cost-benefit ratio, an internal data quality network should be established to improve internal quality control and increase the transparency and comparability of statistics.
- Transparency in the methodologies and procedures used by the statistical offices should be improved through appropriate documentation measures.
- Customer satisfaction surveys should be conducted regularly.
- A concept should be developed for measuring errors and error sources from administrative sources.
- Quality reports should be systematically evaluated for information value and standardized.

We fully support these recommendations but will not consider them further in the final section of this contribution.

The peer review did not deal with some of the problems of the information infrastructure for empirical economic and social research on income, wealth, and poverty. Obviously, they were not within its focus.

### *3.2 Additional problems of the German information infrastructure with respect to scientific research on income, wealth, and poverty*

Statistics that are relevant for research on problems of income, wealth, and poverty in Germany are produced by different public bodies: the German Statistical Offices, the German Federal Employment Agency, the Deutsche Bundesbank, the German Pension Insurance and other social security institutions, and some ministries. Either these statistics are collected in accordance with special statistics laws, or they are produced as part of an institution's general administrative activities, or they are based on surveys carried out by private market research companies and financed by public funds. Regular evaluation processes for these other data collection activities similar to the peer review based on the *European Statistics Code of Practice* for the Federal Statistical Office do not exist for all public bodies. This lack of systematic control casts doubts on the reliability of the various datasets.

Up to now, it has not been possible to draw a comprehensive picture of the distribution of income, wealth, and income poverty within the permanent population of Germany. The German Microcensus, which is compulsory and based on a random sample, covers in principle the entire resident population

but does not contain sufficient information on income and wealth. It therefore only provides the basis for rough analyses of income and poverty.

The EVS contains detailed information on income, wealth, consumption, and savings, but does not cover some population groups, particularly households with very high incomes, persons living in institutions,<sup>16</sup> and the homeless. Additionally, persons with a migration background – especially if they immigrated recently – are grossly underrepresented. One reason for these gaps may be that the EVS is a voluntary survey based on a quota sample instead of a random selection of interviewees. It also uses German questionnaires sent by mail instead of multilingual questionnaires distributed by interviewers. These gaps cause biased results with respect to the distribution of income, wealth, and income poverty. Moreover, because the EVS is a quota sample, confidence intervals cannot be calculated. Although it is not possible to analyze annual income mobility using cross-sectional surveys carried out at five-year intervals, at least one can conduct pseudo-longitudinal analyses based on several of these surveys.<sup>17</sup>

The German contribution to the European Statistics on Income and Living Conditions (EU-SILC) also contains income sources and some indicators of the quality of life, but it neglects components of wealth, consumption, and savings. It is, therefore, only suitable for the analysis of income distributions and income poverty. Each year, one-fourth of the interviewees are selected at random. In its final stage, this survey will be a rotating panel with each interviewee participating for four consecutive years. This will make it possible to analyze annual income mobility and periods of income poverty lasting longer than one year. There are doubts, however, whether the basis for the random selection of interviewees – the so-called “access panel” – is itself a random representation of the resident population of Germany. Moreover, the German contribution to EU-SILC is a voluntary mail survey with questionnaires solely in German, meaning that various groups are underrepresented and that the results will be biased.<sup>18,19</sup>

The data sources are based on different income and wealth concepts: current quarterly household net and gross income, annual household gross

---

16 The concept of institutions is relatively broad. It includes, for example, hospitals, homes for the elderly, nursing homes and orphanages, homes for workers, barracks and prisons, homes for asylum-seekers, monasteries, and similar collective households.

17 See Hauser, R. and Stein, H. (2006): Inequality of the distribution of personal wealth in Germany 1973-98. In: Wolff, E.N. (Ed.) (2006): *International Perspectives on Household Wealth*, Cheltenham/Northampton, 195-224.

18 See Hauser, R. (2007): Problems of the German Contribution to EU-SILC – A research perspective, comparing EU-SILC, Micro census and SOEP. RatSWD Working Paper No. 20.

19 An extensive methodological discussion of the problems of EU-SILC in all Member States of the EU can be found in: European Commission (Ed.) (2007): *Comparative EU statistics on Income and Living Conditions: Issues and Challenges*, Proceedings of the EU-SILC conference (Helsinki, 6-8 November 2006), Methodological and working papers.

and net income from the previous year,<sup>20</sup> monthly gross income from earnings, different lists of components of household wealth, and so on. Although there may be good reasons for different definitions of the variable “income,” it is difficult to combine information from different statistical sources. Moreover, the gross sums of the various kinds of income do not fully correspond with the same categories in the national accounts. For some income categories, these differences amount to more than 30 percent.<sup>21</sup> The differences are even greater with some wealth categories, especially with financial assets as compared to the sums published by the Deutsche Bundesbank. A second serious gap in the statistics on wealth is the neglect of the value of ownership rights in unincorporated businesses. Additionally, the value of consumer durables and cars, antiques, jewelry, and the private ownership of precious metals is not recorded. The distribution of wealth of the resident population should include all assets, irrespective of whether estates and private businesses are located in Germany or abroad. The available household statistics, however, exclude assets located abroad that are not traded on the stock exchange. These various problems result in a considerable underestimation of inequality in the distributions of net equivalent income and net household wealth. To explain the distribution of net wealth, it would be extremely helpful to know the value of the inheritances and gifts *inter vivos* accrued up to the time of interview. Unfortunately, the Income and Consumption Surveys record this information only partially.

The IAB Employment Sample (IABS) contains administrative data on gross labor income of workers and employees, but only up to the limit for social security contributions. For those with higher incomes, the value contained is simply this income ceiling. It is therefore impossible to investigate the upper tail of the labor income distribution or to construct a complete distribution of income from labor.

Minimum benefit regulations, in principle, cover the entire population in the case of net household income below a certain threshold. More than ten percent of the resident population of Germany receives subsidies of various kinds, with standard minimum benefit coverage levels for the entire population. From a social policy point of view, it seems urgent to analyze this large segment of the population below or near the poverty line who have to rely on means-tested minimum benefits. There exist administrative statistics on recipients of minimum benefits under the Unemployment Assistance Law

20 This method is problematic because it only records previous income of persons who are still members of the household at the time of interview.

21 See Sachverständigenrat zur Begutachtung der gesamtwirtschaftlichen Entwicklung, Jahresgutachten 1998/99, Bundestagsdrucksache 14/73 table 57; and Hauser, R. and Becker, I. (2001): Einkommensverteilung im Querschnitt und im Zeitverlauf 1973-1998. Bonn. Table 4.1; and Westerheide, P./Ammermüller, A. and Weber, A. (2005): Die Entwicklung und Verteilung des Vermögens privater Haushalte unter besonderer Berücksichtigung des Produktivvermögens. Bonn. Table 9.

(*Arbeitslosengeld II* as defined in *Sozialgesetzbuch II*, second book of the German social code) and the Social Assistance Law (*Bedarfsorientierte Grundsicherung im Alter und Erwerbsminderung sowie Sozialhilfe* according to *Sozialgesetzbuch XII*). A unified Scientific Use File of minimum income recipients, however, is not available, although this would be of utmost interest for research.<sup>22,23</sup> Only a panel of a sample of the long-term unemployed will become available (PASS as mentioned above). Even the Scientific Use File of former social assistance recipients is no longer available because of changes enacted in 2003 and 2005 that altered the legal framework.

#### 4. Recommendations

To remedy the problems described above, we offer several recommendations for improving the information infrastructure for research on income, wealth, and poverty in addition to the aforementioned recommendations of the Peer Review Group.

- To improve the possibility for in-depth research on recipients of minimum benefits, all statistics on minimum benefits should be harmonized and a single Scientific Use File should be created, similar to the former social assistance file.
- To improve the possibility for in-depth research on household wealth, missing elements of wealth, such as the value of ownership of unincorporated businesses and estates in foreign countries, should be included in the Income and Consumption Survey. It would also be very helpful if a question on the value of all inheritances ever received were included. Additionally, an accurate representation of the foreign population should be guaranteed so that special studies on foreigners living in Germany can be undertaken.
- To simplify work with Scientific Use Files, high-quality reports on all household statistics, including ex post checks based on the Microcensus

---

22 See Hauser, R. (2008): Mindestsicherungsleistungen in Deutschland – ein Plädoyer für eine harmonisierte Gesamtstatistik. In: Rolf, G./Zwick, M. and Wagner, G.G. (Eds.): Fortschritte der informationellen Infrastruktur in Deutschland. Festschrift für Johann Hahlen zum 65. Geburtstag und Hans Jürgen Krupp zum 75. Geburtstag. Baden-Baden, 359-368.

23 In 2008 the Statistical Offices published tables for the recipients of the various minimum benefits and for the recipients of other means-tested benefits. This publication underscores the necessity of harmonizing the concepts and making a Scientific Use File available for all the recipients of the various minimum benefits. See Statistische Ämter des Bundes und der Länder, Soziale Mindestsicherung, Wiesbaden 2008.



and various administrative statistics, should be made available to the scientific community. These reports should also show the differences between survey results and the national accounts and national balance sheets, and possible causes. Moreover, the reliability of the income data reported by the Microcensus should be evaluated.

- To improve the data available for research on income, wealth, and poverty, methods should be developed to combine household surveys with administrative statistics – especially tax statistics – while safeguarding this data for confidentiality. This could be done e.g., by statistical matching.
- To cover the entire resident population, new statistics on the homeless and persons in institutions should be developed, at least including standard demographic variables. Additionally, statistics should be collected on all the institutions in which individuals live.
- The German data from EU-SILC that are handed over to Eurostat should be made available to German researchers through the Research Data Centers. The anonymization undertaken by Eurostat should be considered sufficient for compliance with German data protection regulations since the entire dataset can be obtained for all countries in this form from Eurostat, but at a considerable cost.
- It should be checked whether methodological improvements to the German contribution to EU-SILC can be made. This is all the more important since it will become the main statistical source for the German Poverty and Wealth Reports and the National Action Plans for Social Inclusion. In the long run, this extremely important dataset should be improved by using truly random samples, five waves for each quarter of the rotating panel with the first wave only being used as a pretest, face-to-face interviews with multilingual questionnaires, guaranteeing sole responsibility of the German Federal Statistical Office, and outsourcing fieldwork to a private market research company with a well-trained and permanent staff of interviewers.
- The IAB Employment Sample (IABS) should be expanded to include more precise information on labor income above the limit for social security contributions. This could be done by extending the obligation of employers to report labor income of employees above this limit at least by income brackets.
- It should be guaranteed that the interviewees of all future household surveys financed by public money – especially the ASID (*Alterssicherung in Deutschland*) and AVID (*Altersvorsorge in Deutschland*) – give permission in advance to be included in Scientific Use Files derived from

these surveys, in compliance with data protection laws. These datasets should either be made available for scientific research by transmitting them to the Central Archive in Cologne or by producing and distributing Scientific Use Files through one of the Research Data Centers.

- The problem that legal regulations forbid transmitting Scientific Use Files containing German data to reliable foreign institutions (e.g., universities) should be solved following the suggestions in the 2001 KVI report.<sup>24</sup>
- To improve the quality of statistics produced by public bodies outside the German Federal Statistical Office, a code similar to the *European Statistics Code of Practice* should be developed. Additionally, a regular review process for these other bodies should be introduced, especially for those statistics published by the German Federal Statistical Office but provided by other institutions.

---

24 Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik (KVI) (Eds.) (2001): Wege zu einer besseren informationellen Infrastruktur. Baden-Baden, 152-154.

## 5.3 Family Research

*Johannes Huinink*

Contact:

Johannes Huinink  
University of Bremen  
FB 08 / EMPAS  
FVG-Mitte, Celsiusstrasse  
28359 Bremen  
Germany  
e-mail: [huinink\[at\]empas.uni-bremen.de](mailto:huinink[at]empas.uni-bremen.de)

## Abstract

Family research has become increasingly important in recent years, as reflected in the high public interest in family issues. A number of improvements have been made with regard to the provision of family related data in response to the 2001 KVI report. Family research has profited from these considerably. However, progress in data provision since the beginning of the 21<sup>st</sup> century has been limited, both in terms of the quality and quantity of data produced. In particular, there is still an urgent need for longitudinal data on social and family-related processes dealing with different levels and dimensions of family development. Data are needed not only to describe family change adequately but also to model the determinants and “outcomes” of couple and family dynamics or family relationships over time. What is needed most at present – aside from an improved family data reporting system provided by the official statistical agencies – are panel studies collecting longitudinal (socio-)structural and socio-psychological “on time” information on the dynamics of individuals’ living arrangements over time.

Keywords: family research, longitudinal data, family dynamics

## 1. Introduction

Family research has gained considerable attention in recent years, and the need for more, and more precise, information on various aspects of family dynamics is urgent. This is reflected in the current high public interest in family issues of various kinds: not only family demography in the narrower sense (living arrangements, nuptiality and divorce, fertility) but also aspects of the internal dynamics of close relationships in unions and families (quality and benefits of intimate relationships, parenting, intergenerational relationships, effects of poverty, intra-family violence). These topics often lie on the borderline between social, economic, and psychological research.

Despite the high public interest, family research was not addressed in the first round of KVI advisory reports in 2001 explicitly. In KVI reloaded, aside from the reports on population (Kreyenfeld and Scholz) and intergenerational relationships (Nauck and Steinbach), only this report deals with family research. Family-related issues are still not well represented in a broad range of research fields.

In 2001 the KVI report made several recommendations of particular relevance for the field of family research. These include:

- conducting a census;
- providing long-term institutional support for the German Socio-Economic Panel (SOEP, *Sozio-oekonomisches Panel*);

- allocating more support for prospective and retrospective cohort studies to allow longitudinal analyses of individual development and life courses;
- continuing the German General Social Survey (ALLBUS, *Allgemeine Bevölkerungsumfrage der Sozialwissenschaften*) and the International Social Survey Programme (ISSP);
- improving access to aggregated as well as individual-level data, providing scientific use microdata files, and establishing Research Data Centers; and
- creating opportunities to link data from different data sources.

In nearly all of these areas, improvements made since 2001 have benefited family research considerably. Particular progress has been achieved in access to large datasets (Microcensus) of the official statistical agencies, allowing more valid and detailed description and analysis of changing demographic family structures over time. A census has not materialized up to now, but is in the planning phase. The body of family related panel data allowing longitudinal studies is growing. Family issues have also been addressed in more detail in recent ALLBUS, ISSP, and European Social Survey (ESS) surveys. However, in terms of sample sizes and content, progress since the beginning of the 21<sup>st</sup> century has been limited. It will be argued below that the longitudinal data needed to study social and family-related processes on different levels of family development is still not adequate. The newly launched Panel Analysis of Intimate Relationships and Family Dynamics (pairfam) is one step in overcoming this lack of data in Germany.

## 2. Research in family science

### 2.1 Main research fields

Family research is multidisciplinary by definition, spanning the disciplines of demography, psychology, sociology, economics, anthropology, education, political science, and law. In the following, the main research fields will be addressed systematically from an analytical point of view (Huinink 2008).

*Family and social structure* (the macro perspective): This research field deals with demographic and socio-structural changes in family and living arrangements as well as their structural and institutional embeddedness in our functionally differentiated society. Subfields of research are:

- demography of the family and family types;
- social structure and social inequality of families;
- family as a social institution in welfare states; and

- family and subsystems of the society: demands and achievements of the family related to other subsystems of society.

*Family as a social group* (the meso perspective): This field of research looks at the dynamics of social relationships in private households and families of different kinds and during different phases of family development. Subfields of research are:

- social interaction in couples and families;
- household production and organization of everyday life in couples and families;
- socialization, parenting, and parental transmission; and
- intergenerational relationships.

*Family development over the life course* (the micro perspective): This research field addresses the behavior of individual actors and their motivational structure connected with family development as an interdependent part of the individual life course. Subfields of research are:

- mating, establishing partnerships, family formation and extension;
- stability and disintegration of couple and family relationships; and
- family life and its effects on other domains of the individual life course.

## 2.2 Development in theory and methodological challenges

Investigations in these fields of family research are connected to different theoretical approaches, each requiring different data for empirical investigation and different methods of data analysis. Even a brief overview of the theoretical developments and methodological challenges accompanying them, as well as a review of progress in methods of data analysis, very clearly shows what kind of data are needed to make further progress in family research. Theoretically and methodologically, family research in the social sciences has made considerable progress by overcoming cross-sectional concepts and implementing longitudinal approaches of theoretical and empirical analysis. Family research has profited from new strategies of data collection, especially panel and retrospective survey designs (Mayer 2000; Seltzer et al. 2005). Refined methods of panel and event history analysis allow consideration of different levels of analysis and different dimensions of the life course in the study of couple and family dynamics (Blossfeld and Rohwer 2002; Wu 2004; Singer and Willett 2003; Halaby 2004).

There already exists a broad *theoretical framework* for family research that is still being expanded. The main theoretical paradigm follows a multi-level life-course approach of individual welfare production over time. The

rationale for welfare production can be based on different versions of a theory of individual action over the life course. It makes assumptions about the interdependency among individual action, its contextual conditions on different levels of social processes, and the various closely interrelated dimensions of the individual life course – of which family life is one (Feldhaus and Huinink 2008).

On the macro level, research focuses on social change in the family's structural and institutional context in society. Cohort analysis makes it possible to distinguish period, age, and cohort effects. Research on the meso level addresses the impact of the medium-range social context, the local infrastructure (e.g., child care systems), social networks, working conditions, neighborhoods, etc. On the micro-level of family units and couples, family research examines dynamics of interpersonal relationships of different kinds over time. Here, individual family-related behavior is examined in context, since it is embedded in diverse strata of constantly changing situational conditions.

These fields of family research require specific kinds of data and methodologies. Particularly, family research faces the following *methodological challenges*:

*Third variable phenomenon*: The question of spurious correlation is particularly critical in longitudinal research, especially in self-referential or path-dependent processes like the life course of individuals. One theoretical approach in family research addressing this phenomenon is Hakim's preference theory. It assumes that much of the relationship between family activities and work in later life is preconditioned by early adopted attitudes (Hakim 2000).

*Selection and adaptation*: Processes of selection and adaptation over the life course have to be considered (Lesthaeghe 2002). Intentions, values, aspirations, and frames of action (socio-psychological indicators), shape individual behavior (selection). At the same time new biographical statuses and life course decisions affect individuals' values, attitudes, and aspirations (adaptation). Also individuals' social networks are formed by processes of selection and adaptation.

*Substitution and complementarity*: The life course is a multidimensional process, but little is known about relations of substitution and support between different dimensions of welfare production in different life domains such as family and work.

*Anticipation*: Social actors learn from the past and are restricted in their degrees of freedom for action by past decisions and past behavior. Knowing this, they anticipate future consequences as well as expected changes in the conditions of their action. Future life-course transitions or the "shadow of the future" therefore take on increasing importance in decisions on current behavior (Nauck 2001).



*Couple perspective:* Individual-level family research in sociology is still strongly focused on individual actors and often fails to integrate the perspectives of partners and family members (Lyons and Sayer 2005).

*Cultural comparison:* A great deal of international diversity in family dynamics is due to cultural differences, which are often rooted in processes that took place centuries ago. An example demonstrating the impact of cultural differences that were emerging only decades ago are the different patterns of family development and parental living arrangements in East and West Germany. Up to now, the crucial cultural parameters have not been clearly identified empirically even though we know that socio-structural differences between the populations of the “two Germanys” do not fully explain the divergent behavioral patterns.

Still, there are blind spots in our understanding of the complex individual decision making processes that take place over the life course. In particular, we observe an evident lack of interdisciplinary theory integrating demographic, economic, sociological, and psychological approaches, and a lack of adequate longitudinal data for empirical analysis.

### **3. Status quo: Databases and access**

Although considerable improvements have been made in data provision – with respect to the requirements mentioned above – there are still severe deficits to be noted.

Following the recommendations of the 2001 KVI report, great progress has been made in structural macro and microdata for the demographic analysis of family dynamics. This is thanks to improvements in access to data from official statistical agencies, and in the provision of family-related data from social surveys in the national and international context. However, most of the data available are cross-sectional data enriched by retrospective information. Non-structural information (socio-psychological indicators) is usually only available from cross-sectional surveys or panel studies with long gaps between the few panel waves, while longitudinal data of this type is still lacking.

#### *3.1 Official statistics*

Data from official statistical agencies is useful primarily for descriptive purposes, for example, in reporting changes in family structure. But increasingly official statistical data are also used to model and investigate family

dynamics analytically, thanks primarily to improved access to the Microcensus data.

Statistics on marriage, divorce, and fertility are available but can only be used for descriptive purposes. Up to the year 2007, parity-specific birth statistics cannot be calculated on the basis of data from vital statistics, and the proportion of childless men or women in a particular cohort cannot be estimated accurately.

Major progress has been made in regard to the use of Microcensus data in family research, especially in family demography. Not only are the Microcensus data being used to an increasing degree for descriptive purposes, they are also used more and more in highly differentiated and sophisticated statistical models of family formation and development (e.g., Duschek and Wirth 2005; Kreyenfeld 2001; Kreyenfeld and Geißler 2006; Lengerer and Klein 2007; Lengerer et al. 2007; Wirth 2007; see also the report on demography in this publication).

Up to now, however, users have to struggle with a number of shortcomings. One widely discussed shortcoming is that the Microcensus only considers children of respondents who live in the same household. And surprisingly, questions on day care provisions were not covered in the last Microcensus Act of 2005.

Another problem with using Microcensus panel data is that respondents who change their place of residence drop out of the sample. This means that the panel subsample becomes more and more selective because mobile respondents are underrepresented. If migration behavior is correlated with a dependent variable of interest, biased results can be expected. Nonetheless, it has been investigated whether Microcensus panel data can be used for longitudinal analyses. Kreyenfeld et al. (2007) show that selectivity does not seem to be problematic in case of studying family formation.

Some other surveys by the Federal Statistical Office that are useful for family research should be mentioned. The second time use survey provided considerable information for research on household production in families and households. The Sample Survey of Income and Expenditure (EVS, *Einkommens- und Verbrauchsstichprobe*) can be used to study the economic situation of families and households. The social assistance statistics (*Sozialhilfestatistik*) and youth welfare statistics (*Statistik der Jugendhilfe*, various micro datasets on institutions providing services for children and adolescents) are relevant for family research and are available for several years. These opportunities are not yet being utilized extensively in family research.

### 3.2 Survey data

Longitudinal data for family research are currently available from large-scale studies like the SOEP (Wagner et al. 2007), the German Life History Study (GLHS) (Mayer 2000), the Family Survey of the German Youth Institute (DJI, *Deutsches Jugendinstitut*) (Bien and Marbach 2003), and the Cologne study of upper-level secondary students (*Gymnasiastenstudie*) (Meulemann 1995). The part of the SOEP relevant for family issues has been extended considerably over the years. A questionnaire dealing with newborn and very young children (aged two to three and four to six) and the subsample of adolescents (aged 16-17) provide important data in this context.

The SOEP, the Family Survey, the GLHS, and the Cologne *Gymnasiastenstudie* have brought about a considerable shift in the longitudinal analysis of family dynamics. The SOEP and the GLHS focus on socio-structural as well as demographic data and socio-economic issues. However, they do not allow the study of the interrelatedness between psychological and social dynamics and processes of decision-making about family or intimate relationships.

Studies on fertility and family dynamics that go beyond this “structural bias” are primarily cross-sectional surveys such as the Family and Fertility Surveys (FFS) conducted in the early 1990s and used to this day in international comparative studies. The same is true of the ALLBUS, the European Social Survey (ESS), and the German Family Survey. The latter includes a three-wave panel as a subsample but with a lag of six years between the waves. The Generations and Gender Survey (GGS) conducted under the auspices of the United Nations Economic Commission for Europe (UNECE) (Vikat et al. 2007) will also provide panel data. The time interval between the waves is also quite large here (three years). The German partner in this program is the Federal Institute for Population Research (BiB, *Bundesinstitut für Bevölkerungsforschung*). The first wave of the German GGS was conducted in 2005; in 2006 a migration sample was added (Ruckdeschel et al 2007; Ette et al. 2007); and data collection for a second wave of the German sample started in 2008.

The DJI is running several surveys that are relevant for family research. Among the most important are the Youth Surveys addressing the living conditions and social and political orientations of adolescents and young adults (started in 1992; cross-sectional representative surveys of young people in Germany aged 16 to 29 or 12 to 29), and the Children’s Panel, a longitudinal study started in 2001 that provides data on children’s living situations and the impact of living conditions on children’s individual development.

Socio-psychological determinants of couple and family behavior are being studied more and more in prospective surveys. Yet family behavior is still not being studied extensively in regard to social embeddedness (social

networks, kinship networks) or from the perspective of “linked lives” (Elder 1994). Very few studies so far have attempted to use multi-actor designs to obtain original data on attributes of several related persons. An important exception in Germany is the SOEP, which makes it possible to combine information on the different members of a household.

Longitudinal data on living arrangements of the elderly are provided by the Survey of Health, Ageing and Retirement in Europe (SHARE) (Börsch-Supan et al. 2003). Two waves with respondents aged 50 and higher have been collected so far; the third is on the way. Also, the first and the second versions of the German Ageing Survey (DEAS, *Alterssurvey*) should be mentioned, which have contributed substantially to knowledge on this aspect of family life. But these studies have not been designed as panel studies and are limited to analyzing the dynamics of elderly people’s family lives (Kohli and Szydlick 2000; Tesch-Römer et al. 2006).

To summarize: nearly all fields of family research are making use of information provided by large-scale datasets, and the research infrastructure has improved considerably. However, the richness of the data is often still quite limited due to the theoretical and methodological challenges referred to above.

### *3.3 International perspectives*

An adequate overview of the international situation is impossible to provide in the limited space of this report. Some of the aforementioned German surveys have international counterparts (like the SOEP) or are part of international programs. This is true, for example, of the GGS, the ESS, and the SHARE project.

One international prospective longitudinal study that focuses on family issues and meets the conditions listed above is the Netherlands Kinship Panel Study. It is conducted by the Netherlands Interdisciplinary Demographic Institute in cooperation with several Dutch universities (NKPS; Dijkstra et al. 2004). The research questions focus on issues of intergenerational relationships and solidarity in kinship systems. Two waves of extensive face-to-face interviews have been conducted so far (Wave 1 in 2002-2004, Wave 2 in 2006-2007).

In Britain, important longitudinal data sources (besides the BHPS<sup>1</sup>) are provided by the National Child Development Study and the 1970 British Cohort Study, followed by the National Child Development Study and Millennium Cohort Study. These studies are run by the Centre for Longitudinal Studies UK. The Millennium Cohort Study (MCS) is the UK’s latest longitudinal birth cohort study and follows the lives of a sample of babies

---

1 British Household Panel Study.

born 2000 to 2002. The studies collect information on education and employment, family and parenting, physical and mental health, and social attitudes of large numbers of respondents of selected birth cohorts (Ferri et al. 2003; Dex and Joshi 2005).

## **4. Future developments: Data provision and data access**

### *4.1 Requirements in regard to data*

Although considerable theoretical and methodological progress has been made, some aspects of family research seem to have reached an impasse that crucially needs to be overcome. This is primarily a problem of data, not of theory. Data are needed not only to adequately describe family changes but also to measure the structural and especially socio-psychological determinants and “outcomes” of couples’ and family dynamics and family relationships over time – both retrospectively and prospectively. The resulting data requirements are obvious:

- (1) Data on all levels of analysis are needed. Macro-level data mainly provide information about demographic trends in changing living arrangements, family developments, and social structure. However, these data are also indispensable to multilevel analyses of family processes. We need information about societal conditions of family life (macro-economic, political, and cultural conditions). On the meso-level, what is needed most is information on the regional circumstances (opportunities and restrictions) of parenting, family life, and intergenerational support (support systems and institutions for various needs of couples and families at all stages of their development, labor market conditions, information about company support to families, programs providing childcare for working parents, etc.) as well as data dealing with the social context, kinship structure, and social networks. On the individual level, data are needed to model individual decision processes over time. Moreover, it will be increasingly important to be able to combine data from different sources and different levels of analysis.
- (2) Not only do we need (socio-)structural information (demographic variables; the indicators of “standard demography” as provided by the ALLBUS) on the macro- and micro-levels, but also data on cultural factors impacting family development. Measurement of these factors must be improved by developing instruments to study national or regional cultural patterns in ideas about family or religiosity. The attitudes and socio-psychological dispositions of individual actors will have to be considered in order to model multilevel decision processes

over time, to test bridge hypotheses, and to include the developmental dimension.

- (3) Panel data are needed to test multilevel, dynamic theories of couples' and family development. Cross-sectional data only serve descriptive purposes and help to obtain correct correlations. Data from retrospective surveys are not sufficient, because they can only provide structural information ("life histories") of sufficient validity. At least panel data are needed providing information on life events and socio-psychological dispositions over time. Here, it may even be necessary to implement event-based sampling strategies. Only prospective methods of data collection deliver valid information on social-psychological indicators.
- (4) Because we are dealing with intimate social relationship decisions, individuals (partners, parents and children, grandparents and parents, siblings, etc.) have a strong impact on each other's behavior. Therefore, a multi-actor design is often necessary. In particular, it is virtually impossible to obtain valid proxy information on socio-psychological attributes (such as attitudes) of another person, whether the person reporting is a friend, member of the social network, partner, or parent of the individual being reported on.

To summarize, we need longitudinal, "on time" data from all levels of analysis. At the micro-level, we need more socio-psychological information and multi-actor designs, and high-quality data enabling a differentiated description of family dynamics and changing living arrangements in Germany.

#### 4.2 Official statistics

Many improvements have been made in data from official statistical sources to implement a descriptive reporting system in the field of family research (see the report by Kreyenfeld and Scholz). The quality of the vital statistics of the Federal Republic of Germany on family issues has improved since 2008: they now provide the opportunity to study fertility in line with international standards. The same is true of the Microcensus: at least now, female respondents are asked about their total number of children. This is a very small improvement, however.

Here, more effort should go into using other micro datasets that are relevant for family research, such as the *Statistik der Jugendhilfe* and the like.

### 4.3 Surveys

The empirical basis for descriptive analyses and analytical models in different fields of family research has to be strengthened. Surveys like the Family Survey of the German Youth Institute should therefore be continued: they serve both purposes, as past experiences show. Furthermore, they are essential to a reporting system on family issues, which cannot be created solely using data from official statistical sources (Engstler and Menning 2003).

The other longitudinal research and survey programs mentioned above also have to be continued. A major contribution has been made by the newly launched pairfam with a yearly data collection schedule. Members of three age cohorts (15-17, 25-27, and 35-37 years old in the first wave) will be followed up over subsequent years. The study will provide longitudinal data on the basis of a multi-actor design. Additionally, the partners, parents, and children of the anchor persons will be interviewed. The questionnaires include detailed structural and non-structural information. Particular emphasis is placed on psychological and sociological instruments to obtain prospective information on determinants of establishing intimate relationships and their stability over time; the timing and spacing of fertility; inter-generational relationships and parenting; and social networks. It is important to continue this panel on a long-term basis.

Internationally, the Netherlands Kinship Panel Study will be continued. Also a third wave of the GGS is planned in different European countries.

## 5. Conclusions and recommendations

First of all, there is good reason to demand more attention to the various issues of family research from the German Data Forum (RatSWD) in its efforts to improve the data infrastructure for the social sciences. This need should not be addressed strictly from a demographic perspective – as important as this perspective is. Major aims of such an effort should include:

- (1) Continuing initiatives to improve the family reporting system of the official statistical agencies, allowing for refined description and analysis of family structure and changing living arrangements in Germany. It should be consistent with respective reporting systems in other European countries. The probability of success would increase if scholars from different disciplines of family science in Germany undertook a coordinated initiative, potentially under the auspices of the German Data Forum (RatSWD).

- (2) Providing opportunities to combine data from different sources, allowing the use of more refined models of multilevel analysis in family research. One possibility would be to combine individual-level information of different origins (register data of different kinds) with data on the local family-related infrastructure (day care provision) and data on workplace benefits supporting family needs (working hours policies).
- (3) Continuing and optimizing prospective longitudinal studies collecting structural and socio-psychological information on the dynamics of individuals' living arrangements over time. Because of the special importance of longitudinal research for social research in general, major German panel studies such as pairfam should be integrated into a panel infrastructure covering different fields of social research. Family research would benefit from this considerably.
- (4) Developing new instruments and methods of data collection going beyond the strict panel design with equidistant waves and testing methods of event-based sampling (see the report by Riediger).
- (5) Improving the conditions for comparative longitudinal research by pursuing closer international cooperation.



## References:

- Bien, W. and Marbach, J.H. (Eds.) (2003): Partnerschaft und Familiengründung – Ergebnisse der dritten Welle des Familien-Survey. Opladen.
- Blossfeld, H.-P. and Rohwer, G. (2002): Techniques of event history modeling. New approaches to causal analysis. Second Edition. Hillsdale (NJ),
- Börsch-Supan, A./Jürges, H. and Lipps, O. (2003): SHARE - Measuring the aging process in Europe. ZUMA Nachrichten 53, 96-113.
- Dex, S. and Joshi, H. (Eds.) (2005): Children of the 21st century: From birth to nine months. Bristol.
- Dijkstra, P.A. et al. (2004): Codebook of the Netherlands Kinship Panel Study, a multi-actor, multi-method panel study on solidarity in family relationships. The Hague.
- Duschek, K.-J. and Wirth, H. (2005): Kinderlosigkeit von Frauen im Spiegel des Mikrozensus. Eine Kohortenanalyse der Mikrozensus 1987 bis 2003. Wirtschaft und Statistik 8/2005, 800-820.
- Elder, G.H. Jr. (1994): Time, Human Agency, and Social Change: Perspectives on the Life Course. Social Psychology Quarterly 57, 4-15.
- Engstler, H. and Menning, S. (2003): Die Familie im Spiegel der amtlichen Statistik. Berlin.
- Ette, A./Hullen, G./Leven, I. and Ruckdeschel, K. (2007): Generations and Gender Survey. Dokumentation der Befragung von türkischen Migranten in Deutschland. Materialien zur Bevölkerungswissenschaft, Heft 121b. Wiesbaden.
- Feldhaus, M. and Huinink, J. (2008): Family Research from the Life Course Perspective. International Sociology. [Forthcoming].
- Ferri, E./Bynner, J. and Wadsworth, M. (Eds.) (2003): Changing Britain, Changing Lives: Three generations at the turn of the century. London.
- Hakim, C. (2000): Work-Lifestyle Choices in the 21st Century: Preference Theory. Oxford.
- Halaby, Ch. (2004): Panel Models in Sociological Research. Annual Review of Sociology 30, 507-544.
- Huinink, J. (2008): Gegenstand der Familiensoziologie. In: Schneider, N.F. (Ed.): Lehrbuch Moderne Familiensoziologie. Opladen.
- Kohli, M. and Szydlik, M. (Eds.) (2000): Generationen in Familie und Gesellschaft, Opladen.
- Kreyenfeld, M. (2001): Employment and Fertility – East Germany in the 1990s Rostock. [Unpublished dissertation].
- Kreyenfeld, M. and Geisler, E. (2006): Müttererwerbstätigkeit in Ost- und Westdeutschland. Zeitschrift für Familienforschung 18, 333-360.
- Kreyenfeld, M./Zühlke, S. and Konold, M. (2007): Ökonomische Basis der Familiengründung – Analysen mit den Daten des Mikrozensus-Panels. Presentation on the 5. Nutzerkonferenz „Forschungen mit dem Mikrozensus: Analysen zur Sozialstruktur und Arbeitsmarkt“, 15-16 November 2007, Mannheim.
- Lengerer, A./Janßen, A. and Bohr, J. (2007): Familiensoziologische Analysepotenziale des Mikrozensus. Zeitschrift für Familienforschung 19, 186-209.

- Lengerer, A. and Klein, Th. (2007): Der langfristige Wandel partnerschaftlicher Lebensformen im Spiegel des Mikrozensus. *Wirtschaft und Statistik* 4, 433-447.
- Lesthaeghe, R. (Ed.) (2002): *Meaning and Choice: Value Orientation and Life Course Decisions*. NIDI/CBGS Publication. The Hague/Brussels.
- Lyons, K.S. and Sayer, A.G. (2005): Longitudinal Dyad Models in Family Research. *Journal of Marriage and Family* 67, 1048-1060.
- Mayer, K.U. (2000): Promises fulfilled? A review of 20 years of life course research. *Archives européennes de sociologie* 41, 259-282.
- Meulemann, H. (1995): Die Geschichte einer Jugend. Lebenserfolg und Erfolgsdeutung ehemaliger Gymnasiasten zwischen dem 15. und 30. Lebensjahr. Opladen.
- Nauck, B. (2001): Der Wert von Kindern für ihre Eltern. „Value of Children“ als spezielle Handlungstheorie des generativen Verhaltens und von Generationenbeziehungen im interkulturellen Vergleich. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 53, 407-435.
- Ruckdeschel, K./Ette, A./Hullen, G. and Leven, I. (2007): *Generations and Gender Survey. Dokumentation der ersten Welle der Hauptbefragung in Deutschland. Materialien zur Bevölkerungswissenschaft, Heft 121a*. Wiesbaden.
- Seltzer, J.A. et al. (2005): Explaining Family Change and Variation: Challenges for Family Demographers. *Journal of Marriage and Family* 67, 908-925.
- Singer, J.B. and Willett, J.B. (2003): *Applied Longitudinal Data Analysis*. Oxford.
- Tesch-Römer, C./Engstler, H. and Wurm, S. (Eds.) (2006): *Altwerden in Deutschland – Sozialer Wandel und individuelle Entwicklung in der zweiten Lebenshälfte*. Wiesbaden.
- Vikat, A. et al. (2007): *Generations and Gender Survey (GGS) Towards a Better Understanding of Relationships and Processes in the Life Course*. *Demographic Research*, 17, pp. 389-440
- Wagner, G.G./Frick, J.R. and Schupp, J. (2007): The German Socio-Economic Panel Study (SOEP) – Scope, Evolution and Enhancements. *Schmollers Jahrbuch* 127 (1), 139-169.
- Wirth, H. (2007): Kinderlosigkeit im Paarkontext von hochqualifizierten Frauen und Männern – Eine Folge von Bildungshomogamie? In: Konietzka, D. and Kreyenfeld, M. (Eds.): *Ein Leben ohne Kinder? Kinderlosigkeit in Deutschland*. Wiesbaden.
- Wu, L.L. (2004): Event History Models for Life Course Analysis. In: Mortimer, J.T. and Shanahan, M.J. (Eds.): *Handbook of the Life Course*. New York.

## 5.4 Intergenerational Relationships

*Bernhard Nauck and Anja Steinbach*

Contact:

Bernhard Nauck  
Anja Steinbach  
Chemnitz University of Technology  
Department of Sociology  
Thueringer Weg 9  
09107 Chemnitz  
Germany  
e-mail: [bernhard.nauck\[at\]phil.tu-chemnitz.de](mailto:bernhard.nauck[at]phil.tu-chemnitz.de)  
[anja.steinbach\[at\]phil.tu-chemnitz.de](mailto:anja.steinbach[at]phil.tu-chemnitz.de)

## Abstract

Intergenerational relationships within family and kinship structures have become a salient issue in scientific research. The major reasons for this are the intense demographic changes that occurred throughout the twentieth century, such as an increased life expectancy in combination with decreased fertility, and the implications of this for the major institutions of the social welfare state. This has resulted in several larger studies that can serve as the basis for an analysis of the situation as it impacts older people, including the German Socio-Economic Panel (SOEP, *Sozio-oekonomisches Panel*), the Generations and Gender Survey (GGS), the Family Survey, the German Ageing Survey (DEAS, *Deutscher Alterssurvey*), the Survey on Health, Ageing and Retirement in Europe (SHARE), and the Panel Analysis of Intimate Relationships and Family Dynamics (pairfam). However, an overarching theoretical and research perspective on intergenerational relationships from their creation (fertility) through parenting to the longest lasting relationships between adults of different generations is still lacking. In order to overcome this deficiency, this paper recommends that future data structures obtain information on intergenerational relationships through data that is obtained (1) simultaneously and is theoretically complete, (2) in a lifespan perspective, (3) from a panel design and (4) a multi-actor design. Studies should (5) account for cultural variability of intergenerational relationships and (6) for institutional settings in cross-national comparisons.

Keywords: intergenerational relationships, intergenerational solidarity, life course, demographic change, ageing, panel studies

## 1. Introduction

Intergenerational relationships within the family and kinship structures have become a salient issue in public discourse as well as in scientific research. The major reasons for this are the intense demographic changes that occurred throughout the twentieth century, such as an increased life expectancy in combination with decreased fertility and their implications for the major institutions of the social welfare state. Since the end of the 1990s in the social sciences, this has resulted in the planning and realization of several larger studies on the situation of older people in Germany and Europe, including their relationships to family members. It has also resulted in the implementation of instruments measuring parent-child-relationships in already existing or newly initiated longitudinal surveys. Due to the urgent political issues surrounding these questions, initial research on intergenerational relationships in families in the context of demographic changes has focused on relationships between *aged* parents and their *adult* children, focusing gen-

erally on the question of family-based care and intergenerational solidarity in later stages of life. Accordingly, data collection has concentrated on relationships between children and their (very) old parents. The research domain has therefore been located at the intersection of family research and research on aging. Labeled *research on intergenerational relationships*, it is at present clearly distinguished from *research on parent-child relationships* as a classical research domain of parenting within developmental psychology and socialization research. *An overarching theoretical and research perspective on intergenerational relationships from their creation (fertility) through parenting to the longest lasting relationships between adults of different generations is still lacking.*

## **2. Theoretical developments and research questions**

### *2.1 Theoretical developments*

Most literature on intergenerational relationships starts with a reflection on the family-in-crisis hypothesis. To test this hypothesis, but also to give a descriptive picture, various aspects of these relations are considered. The most important contributions to this framework include the theory of intergenerational solidarity (Bengtson and Roberts 1991; Bengtson 2001) and work on ambivalence (Lüscher and Pillemer 1998; Pillemer and Lüscher 2004). These contributions to the discourse consider many different aspects of contact and supportive behavior within the family and between generations. Based heavily on social exchange theory, intergenerational relationships are understood as any form of exchange between generations. Six exchange dimensions are distinguished, namely structural, associative, affective, consensual, normative, and functional solidarity.

The *structural dimension* refers to the opportunity structure that determines the specific ways that family interactions are realized. Typical measurements are geographical distance and residential proximity, but availability of kin, parents, children, and siblings, as well as their age, sex, marital status, health status, and working arrangements are also seen as important factors in structural solidarity. The *associative dimension* refers to the amount and kind of intergenerational contact, either face-to-face or by phone, e-mail, or other means. Therefore, frequency and intensity of contact can be distinguished. The *affective dimension* comprises emotional closeness as well as conflict as measures of the quality of the relationship between children and their parents. The *consensual dimension* measures the degree of agreement in values and beliefs – whatever the specific content of these convictions may be. The *normative dimension* refers to the extent of commitment to filial and parental obligations by the respective members of

intergenerational relationships. The *functional dimension* measures all kinds of financial, instrumental, and emotional support that are exchanged between parents and children.

However, the various types of interaction between generations are not always positive. Intergenerational relations can – and typically do – comprise both positive and negative components, and thus are to some extent *ambivalent*. This is due to the social character of intergenerational relationships, which is in most cases unavoidable and inescapable, rather “diffuse” in their exchange, and thus “packaged.” It is an open debate whether ambivalence should be measured directly, for example by asking about the amount of simultaneously positive and negative (i.e., ambivalent) emotions or whether conclusions about ambivalence should be arrived at indirectly, from the extent of simultaneous emotional closeness and conflict between generations.

For a long time, theoretical discussions around intergenerational relationships were (and to some extent still are) limited to the question of whether these named dimensions are adequate (Szydlik 2000) or complete (Bengtson et al. 2002; Lüscher and Pillemer 1998). More recent discussions have become increasingly critical and point out the theoretical deficits of the well-established paradigms (Dallinger 2002; Hammarström 2005; Katz et al. 2005; Grünendahl and Martin 2005). However, serious attempts to provide theoretical explanations for the emergence of and changes within and between the respective dimensions of intergenerational relationships are still very rare (Merz et al. 2007; Steinbach and Kopp 2008a). In summary, one may state that apart from the heuristic model of Szydlik (2000), which includes the associative, affective, and functional dimensions and relates them to opportunity, need, family, and cultural-contextual structures, there is no elaborated theory of intergenerational relationships.

## 2.2 Research questions

Empirically, several different analytical strategies can be distinguished with regard to the respective aspects of intergenerational relationships (Steinbach and Kopp 2008a).

### 2.2.1 Intergenerational solidarity

Within this research domain, a first group of studies focuses on the internal structure of the dimensions of intergenerational solidarity (Atkinson et al. 1986; Roberts and Bengtson 1990; Rossi and Rossi 1990). A second group of studies tries to combine these different dimensions of intergenerational relations, aiming at the construction of family typologies (Bengtson 2001; Giarrusso et al. 2004; 2005; Katz et al. 2005; Van Gaalen and Dykstra 2006;

Silverstein et al. 1994; Steinbach 2008). A third group of studies deals with the different perspectives that parents and children have on their respective relationships. This research has resulted in the so-called “intergenerational stake hypotheses” (Bengtson and Kuypers 1971), and has recently initiated several subsequent replications (Aquilino 1999; Giarrusso et al. 1995; Trommsdorff and Schwarz 2007). The fourth and largest group of studies can be characterized by varied attempts to identify independent socio-structural, intrafamilial, or intergenerational factors that determine intergenerational relationships. Determinants of the degree of emotional closeness, the frequency of contact, and the level of exchange are of particular interest (Attias-Donfut 2000; Hank 2007; Kaufman and Uhlenberg 1998; Klaus 2009; Kohli et al. 2005; Lawton et al. 1994; Parrott and Bengtson 1999; Roberts and Bengtson 1990; Rossi and Rossi 1990; Spitze and Logan 1991; Steinbach and Kopp 2008b; Szydlik 1995; 2000). But also, for example, conflict (Szydlik 2008), ambivalence (Pillemer and Suitor 2002), and inheritance (Kohli 2004; Lauterbach and Lüscher 1996; Nauck 2009b; Szydlik 1999; 2004; Szydlik and Schupp 2004) are important empirical research subjects. The results of all these studies stand in sharp contrast to the popular perception of weakening ties between generations in “postmodern” families. Instead, intergenerational relationships have become – despite changing demographic structures – increasingly important for family members and are obviously one of the major mechanisms of social integration in functionally differentiated societies.

#### 2.2.2 Gender

One structural variable that has played an important role over the years and thus will be given particular attention at this point is that of gender. Empirical results show consistently that the respective combination of gender across generations structures the relationship considerably; that is, there is a rank order in the closeness of the relationship from mother-daughter to mother-son, father-daughter and father-son relationship (Kaufman and Uhlenberg 1998; Nauck 2009a; Rossi 1993; Szydlik 1995). Women – especially from the older generation – function as “kinkeeper” (Atkinson et al. 1986; Rossi and Rossi 1990), maintaining the relationships and providing support. Moreover, women are prone to find themselves in a “sandwich” situation, with simultaneous care activities for both the older and the younger generation within the family. However, this phenomenon becomes rarer with increasingly healthy aging and extended age differences between generations (Kohli and Künemund 2005a; Künemund 2006).



### 2.2.3 Life course

In recent years, research on intergenerational relationships has also adopted a life-course perspective. Although cross-sectional data are predominantly used, the interesting research question has become whether early life stages have an important impact on the intergenerational relationships in later life. From this perspective, one topic of investigation has been the degree to which parents' early transfers to their young adult children affect the children's propensity in middle age to provide social support to their aging parents (Silverstein et al. 2002) and how life-course transitions experienced by each generation affect the quality of relationships between adult children and their parents (Kaufman and Uhlenberg 1998). In particular, the separation and divorce of parents as a potential obstacle in later life intergenerational relationships has become an important research question (Aquilino 2005; Kalmijn 2008; Lin 2008). Likewise, the relationship between attachment patterns in early childhood and the exchange of support in later life stages has arisen as a significant theme in the research (Cicirelli 1993; Merz et al. 2008; Schwarz and Trommsdorff 2005).

### 2.2.4 Cross-national comparisons

The establishment of cross-national and cross-cultural comparative datasets has made it possible to investigate intergenerational relationships in a comparative perspective. Such research programs, predominantly based on cross-sectional data, exist especially for East Asian societies (Hermalin 2002), for Europe (Albertini et al. 2007; Brandt and Szydlik 2008; Haberkern and Szydlik 2008; Hank 2007; Katz et al. 2005), and to some extent for comparisons across continents (Nauck 2009a; 2009b; Nauck and Suckow 2006; Nauck and Yi 2007; Trommsdorff and Nauck 2005). The predominant focus in cross-national research is the interrelationship between the social-political regimes on the one hand and the structure of intergenerational exchange relationships on the other, for example, whether social-political measures and incentives may deteriorate intergenerational support and solidarity (crowding out) or whether they enable and enhance them (crowding in) (Künemund 2008). Empirical research provides some evidence that economic transfer and care provisions by the welfare state do not edge out intergenerational support – both seem to complement each other (Armi et al. 2008; Attias-Donfut 2000; Brandt and Szydlik 2008; Künemund and Vogel 2006). Moreover, empirical evidence has suggested that social-political regimes and individual involvement in intergenerational support interact strongly (Haberkern and Szydlik 2008).

### 2.2.5 Social and demographic change

Major demographic trends in the twentieth century had a strong impact on the analysis of intergenerational relationships. One emerging research domain is the analysis of intergenerational relationships beyond the parent-child dyad, namely grandparent-grandchildren relationships (Hank and Buber 2009; Harper 2005; Hoff 2007; King and Elder 1995; 1997; Mueller and Elder 2003). Increased life expectancy in welfare societies has not only resulted in the prolonged common lifetime of parents and children, but also in the increased existence of families with three and even four generations (Hoff 2006; Lauterbach 1995; Lauterbach and Klein 2004). This phenomenon, together with the decline of horizontal kinship relationships because of reduced fertility, was coined as the “beanpole family” (Bengtson et al. 1990) and described as a multi-local extended family structure (Bertram 2003; Lauterbach 2004). This development has stimulated questions about the extent to which relationships between generations are interwoven (Friedman et al. 2008) and the extent to which grandparent-grandchildren relationships are comparable to parent-child relationships (Hoff 2007).

Another major demographic trend is the increased number of immigrants and their aging patterns (Dietzel-Papakyriakou 1993; Nauck 2007). Empirical research has been dedicated to the question of whether intergenerational relationships differ between migrant and native families, between immigrant families of different origin and within different receiving contexts, and how these relationships are maintained across national borders (Attias-Donfut and Wolff 2008; Baykara-Krumme 2008a; 2008b; Komter and Schans 2008; Nauck 2001; Nauck and Kohlmann 1998).

## 3. Status quo: Databases and access

Meanwhile, several datasets exist that can be used for the analysis of intergenerational relationships. In Germany, these include on the one hand the large-scale datasets like the SOEP, the GGS, and the Family Survey, which encompass large age brackets. On the other hand, datasets from aging research are also available, such as the study Old Age and Autonomy: The Role of Service Systems and Intergenerational Solidarity (OASIS), DEAS or SHARE, which mostly concentrate on the population from age forty onwards. Additionally, the dataset of pairfam will be available soon, which will provide data on the intergenerational relationships of younger respondents (aged between fifteen and fifty) with their respective parents. GGS, OASIS, and SHARE are cross-national comparative research programs that allow for analyzing the German situation in an international perspective.

Other important international studies of intergenerational relations without German samples are the Netherlands Kinship Panel Study (NKPS) (Dykstra 1999; Dykstra et al. 2006) and two studies from the United States: the Longitudinal Study of Generations (LSOG) (Mangen et al. 1988; Giarrusso and Zucker 2004), and the American National Survey of Families and Households (NSFH) (Sweet and Bumpass 2002). A systematic comparison of the existing datasets on an international level is provided in the appendix. The following brief description concentrates on German datasets and those with German participation.

*German Socio-Economic Panel.* The SOEP of the German Institute for Economic Research (DIW Berlin, *Deutsches Institut für Wirtschaftsforschung*) has collected detailed annual data to measure the stability and change of living conditions in Germany since 1984 (Frick 2007). Since the early 1990s, it has extended its scope to include some instruments on intergenerational relationships, such as residential distance and emotional closeness to biological parents and to the closest living son or daughter (if the respondent has more than one) (1991, 1996, 2001, 2006). Since 1984, the amount of financial transfers between generations is also captured (exception: 1992 and 1994), including intergenerational donations, inheritance, and bequest.

*Generations and Gender Survey.* The GGS is the German version of an international research program in sixteen countries. It is merged into the “Generations and Gender Program” (GGP) of the United Nations Economic Commission of Europe (UNECE) (Ruckdeschel et al., 2006). The first two waves were completed in 2005 and 2008. The GGS contains questions regarding residential distance, frequency of contact and emotional closeness to parents and children, filial obligations, and daytime care of grandchildren. Financial, instrumental, and emotional support is captured with a network generator, within which family members can be named.

*Family Survey.* The Family Survey of the German Youth Institute (DJI, *Deutsches Jugendinstitut*) (Bien and Marbach 2008) includes a three-wave panel as a subsample but with a lag of six years between the waves (1988, 1994, 2000). Emotional closeness and exchange of financial support is captured with a network generator, within which family and kinship members can be named. For all named individuals, information on relationship quality, residential distance, and frequency of contact are available.

*Old Age and Autonomy: The Role of Service Systems and Intergenerational Family Solidarity.* OASIS is conducted in five countries, including Germany (Tesch-Römer et al. 2000; Lowenstein and Ogg 2003). Data collection took place in 2000 in urban regions only. The disproportionate stratified sample starts at twenty-five years old and over-represents individuals older than seventy-five. Intergenerational relationships are measured based on the dimensions residential distance, frequency of contact (to parents

and all children), emotional closeness, conflict and ambivalence, consensus (degree of similarity on opinions and values between parents and focus child) and the agreement on filial obligations. Mutual support is captured by data on financial, emotional, and instrumental help within the last twelve months. Moreover, grandparent-grandchildren relationships are covered with regard to residential distance, frequency of contact, and support.

*German Ageing Survey.* The DEAS of the German Center of Gerontology is a study of the living situation of people aged forty and older in Germany. Three waves were completed in 1996, 2002, and 2008 (Kohli and Kunemund 2005b; Tesch-Römer et al. 2002; 2006). For all children and for individuals with whom the respondent predominantly grew up, and for up to eight additional network members, the following dimensions of intergenerational relations are captured: residential distance, frequency of contact, and emotional closeness. Exchange of support is part of a network generator, within which up to five persons may be named with whom the respondent exchanges financial, instrumental, and emotional support. Daytime care of grandchildren is also captured, as well as inheritance and bequest.

*Survey of Health, Ageing and Retirement in Europe.* SHARE is an international longitudinal research program and comprises fifteen countries in Europe (Bösch-Supan and Jürges 2005; Bösch-Supan et al. 2005) with three panel waves in 2004-2005, 2006-2007, and 2008-2009. The first wave captured target persons of fifty years and older and their household partners. Intergenerational relationships are covered by the dimensions residential distance, frequency of contact, and emotional closeness to parents and all children living outside the respondent's household. Received help was captured by questions surveying sources of material and financial transfer within the last twelve months and whether caring, if necessary, was received. Up to three individuals can be named. Daytime care of grandchildren is also captured.

*Panel Analysis of Intimate Relationships and Family Dynamics.* pairfam is a comprehensive research program about partnership and family development in Germany (Huinink and Feldhaus 2008). It is based on a cohort design, comprising three cohorts of 15 to 17, 25 to 27, and 35 to 37 years-old target persons respectively in combination with a multi-actor design, including the respective partner, both parents, and children of eight years and older. Data collection of the first wave takes place in 2008-2009, with thirteen further waves currently planned. In the first wave, short versions of instruments on intergenerational relationships are applied (residential distance, frequency of contact, and emotional closeness). From the second wave onwards, comprehensive instruments on residential distance, frequency of contact, emotional closeness, conflict, ambivalence, and agreement to filial obligations, as well as material, financial, instrumental, and emotional transfers will be used, targeting the relationship to both biological parents and, if rele-

vant, stepparents. The multi-actor design implies that from the second wave onwards, (step)parents will provide information about their perspective on the intergenerational relationship towards the target person and his or her respective partner, and that the partner will provide information on his or her relationship to the parents-in-law.

Although surveys that include topics around intergenerational relationships have grown considerably during the last decade, there are still obvious data deficits – especially in Germany:

- As the data on intergenerational relationships are in most cases limited to the measurement of selective dimensions of intergenerational solidarity, their internal structure and mutual influences can not be tested.
- Most studies originate in the field of social gerontology with a focus on the elderly, their family support, and its relation to institutional care-taking.
- Most studies are highly selective in the choice of the studied intergenerational relationships, such that only the relationship to the emotionally and geographically closest child or parent is surveyed, resulting in a positive bias in the scientific description of intergenerational relationships.

For a better understanding of intergenerational relationships in present society, a life-course perspective that covers the development of intergenerational relationships across the entire lifespan and under varying family settings, including non-biological forms of parent-child relations, is needed. Only then can valid measurements of intergenerational solidarity, conflict, and separation in their various dimensions be obtained and thus also allow researchers to make informed estimates about the future potential and development of intergenerational solidarity and social integration.

#### **4. Future developments and recommendations**

The diagnosis of these deficits allows us to suggest some recommendations for creating the data structures necessary for future research.

- Data on intergenerational relationships should be *obtained simultaneously and should be theoretically complete*. That is, all dimensions of the well-established model of intergenerational solidarity and its extensions into conflict have to be measured. Only this will allow for the investigation of the interrelationship between the various dimensions of

intergenerational relations – an issue of significant scientific and practical interest.

- Data on intergenerational relationships should be targeted to a *lifespan perspective*. Intergenerational solidarity in later life stages depends on intergenerational experiences in formative life stages, trajectories, and alternate options and obligations during the previous life course, and is therefore path dependent. The interdependence of generations during the entire lifespan is one of the most important desiderata in this research domain.
- The study of intergenerational relationships needs *panel designs*. Only panel designs allow for the analysis of the creation and the development of intergenerational relationships in specific stages of the life course. They should be complemented by retrospective information on critical life events and related to intergenerational relationships in the past biography of the respondents and his or her family members.
- Methodological research is urgently needed with regard to the *measurement intervals* for intergenerational relationships. Since previous research has concentrated on the most stable and most harmonious relationships in later life, this research provides no knowledge base for an adequate measurement of instable, disruptive, or conflictual parent-child relationships.
- The study of intergenerational relationships should include a *multi-actor design* in order to include the perceptions, evaluations, needs, and resources on both sides of an intergenerational relationship, which is, by nature, asymmetrical and thus prone to differences between members. Moreover, each individual operates and can be statistically modeled within the context of other's actions. Comprehensive analyses of multilevel panel data on intergenerational relationships will be a major research agenda in this realm.
- The study of intergenerational relationships should account for *cultural variability and diversification*. The increasing number of individuals with a migration background is resulting in an increased variability of values related to filial and parental obligations, of arrangements in intergenerational support, and of wealth flows between generations. Thus, specific measurements should be included not only to accommodate migrant and minority situations, but also a range of cross-culturally informed adaptations, which still have to be developed and tested.
- The emergence of multi-local and multi-generational family structures demands special provisions in the collection of data, and, in most cases, a *multi-method-design*. As the study of intergenerational relationships

cannot be based on a standard representative survey design, where all respondents are accessed with the same data collection method, it will be necessary to use a combination of various obtainable methods, such as mail survey, computer-assisted telephone interview (CATI), computer-assisted personal interview (CAPI), computer-assisted self interview (CASI), paper and pencil interview (PAPI), or computer-assisted web interview (CAWI). However, no systematic results are available yet to allow for estimation of the respective advantages and disadvantages of each method within this specific research field.

- For a full understanding of the interplay between institutional settings in the respective social context and the specific structure of intergenerational relationships, *cross-national and cross-cultural comparisons* are needed. To achieve this goal, concepts and measurements have to be standardized and tested for linguistic and functional equivalence. These efforts require a specific infrastructure and extended time for development, both of which are typically disregarded in the funding of comparative research programs. Effective international collaboration needs an additional infrastructure from which standardization and equivalence testing can be coordinated.

The study of intergenerational relationships is an emerging and expanding research domain in the social sciences. It is situated at the intersection of a micro-social level of interactionist family sociology, the meso-level of network analysis and human ecology, and the macro-level of societal integration and social inequality. Its constantly developing nature requires one to adopt a lifespan perspective, which both asks for and permits interdisciplinary cooperation, including a large array of disciplines including developmental psychology, social gerontology, demography, economy, and sociology.

**Appendix: Overview of surveys which include measures of intergenerational relationships**

Study	Full Name	Institution	Data Collection	Unit of Observation
SOEP	German Socio-Economic Panel Study	German Institute for Economic Research, Berlin	Since 1984 IGR: 91, 96, 01, 06	Households (n=10,000) Individuals (n=20,000)
Family Survey	German Family Survey	German Youth Institute, Munich	1988/1990 1994 2000	Individuals (n=10,000)
GGG	Generations and Gender Survey	Federal Institute for Population Research, Wiesbaden	2005 2008 (2011)	Individuals (n=10,000)
OASIS	Old Age and Autonomy: The Role of Service Systems and Intergenerational Family Solidarity	The German Centre of Gerontology, Berlin	2000	Individuals (n=1,300)
DEAS	German Ageing Survey	The German Centre of Gerontology, Berlin	1996 2002 2008	Individuals (n=5,000)
SHARE	Survey of Health, Aging and Retirement in Europe	Mannheim Research Institute for the Economics of Aging, Mannheim	2004/5 2006/7 2008/9	Individuals (n=3,000) Partner
pairfam	Panel Analysis of Intimate Relationships and Family Dynamics	Universities of Bremen, Chemnitz, Mannheim, Munich	2008/9 + 13 waves	Individuals (n=12,000) Partner, Children, Parents



Universe	Countries	Migrant Sample	Dimensions of Intergenerational Relations
18+ Panel	DE	684 Foreigners (1994+)	Distance, Contact, Emotional Closeness, Transfer
18-55 Mixed Design	DE		Distance, Contact, Emotional Closeness, Satisfaction, Transfer
18-79 Mixed Design	DE (AU, BE, BG, CZ, EE, FR, GE, HU, IT, JP, LT, NL, NO, RO, RU)	4,000 Turks (in 2006)	Distance, Contact, Satisfaction, Filial Obligations, Transfer
25+ Cross-sectional	DE (IL, NO, ES, UK)		Distance, Contact, Emotional Closeness, Consensus, Filial Obligations, Transfer
40+ Mixed Design	DE	586 Foreigners (in 2002, 2008)	Distance, Contact, Emotional Closeness, Transfer
50+ Panel	DE (AT, BE, DK, FR, GR, IT, ES, CH, NL/CZ, IE, PL/SI)		Distance, Contact, Emotional Closeness, Transfer
15/25/35 Panel	DE	300 Turks (in 2008)	Distance, Contact, Emotional Closeness, Conflict, Filial Obligations, Transfer

<b>Study</b>	<b>Full Name</b>	<b>Institution</b>	<b>Data Collection</b>	<b>Unit of Observation</b>
NKPS	Netherlands Kinship Panel Study	Netherlands Interdisciplinary Demographic Institute, The Hague; NL	2002/4 2006/7	Individuals (n=10,000) Partner, Children, Parents, Siblings
LSOG	Longitudinal Study of Generations	University of Southern California, Los Angeles, USA	1971, 1985, 1988, 1991, 1994, 1997	Families (n=300) Members of four Generations (in 2000)
NSFH	American National Survey of Families and Households	Center for Demography, University of Wisconsin, USA	1987/8, 1992/4; 2001/2	Individuals (n=13,000) Partner, Children

Universe	Countries	Migrant Sample	Dimensions of Intergenerational Relations
18-79 Panel	NL	1,400 Migrants (in 2002, 2006)	Distance, Contact, Relationship Quality, Conflict, Filial Obligations, Transfer
18+ Panel	US		Distance, Contact, Emotional Closeness, Conflict, Consensus, Filial Obligations, Transfer
18+ Panel	US	Oversample of Blacks/Puerto Ricans/Mexicans	Distance, Contact, Relationship Quality, Transfer

## References:

- Albertini, M./Kohli, M. and Vogel, C. (2007): Intergenerational transfers of time and money in European families: common patterns – different regimes? *Journal of European Social Policy* 17 (4), 319-334.
- Aquilino, W.S. (1999): Two Views on One Relationship: Comparing Parents' and Young Adults Children's Reports of the Quality of Intergenerational Relations. *Journal of Marriage and the Family* 61, 858-870.
- Aquilino, W.S. (2005): Impact of Family Structure on Parental Attitudes toward the Economic Support of Adult Children Over the Transition to Adulthood. *Journal of Family Issues* 26 (2), 143-167.
- Armi, F./Guilley, E. and Lalive D'Epinay, Ch.J. (2008): The Interface between Formal and Informal Support in Advanced Old Age: A Ten-Year Study. *International Journal of Aging and Later Life* 3 (1), 5-19.
- Atkinson, M.P./Kivett, V.R. and Campbell, R.T. (1986): Intergenerational Solidarity: An Examination of a Theoretical Model. *Journal of Gerontology* 41 (3), 408-416.
- Attias-Donfut, C. (2000): Familialer Austausch und soziale Sicherung. In: Kohli, M. and Szydlik, M. (Eds.): *Generationen in Familie und Gesellschaft*. Opladen.
- Attias-Donfut, C. and Wolff, F.-Ch. (2008): Patterns of Intergenerational Transfers Among Immigrants in France: Comparative Perspective. In: Saraceno, Ch. (Ed.): *Families, Aging and Social Policy. Intergenerational Solidarity in European Welfare States*. Cheltenham/Northampton.
- Baykara-Krumme, H. (2008a): *Migrant Families in Germany: Intergenerational Solidarity in Later Life*. Berlin.
- Baykara-Krumme, H. (2008b): Reliable Bonds? A Comparative Perspective of Intergenerational Support Patterns Among Migrant Families in Germany. In: Saraceno, Ch. (Ed.): *Families, Aging and Social Policy. Intergenerational Solidarity in European Welfare States*. Cheltenham/Northampton.
- Bengtson, V.L. (2001): Beyond the Nuclear Family: The Increasing Importance of Multigenerational Bonds. *Journal of Marriage and the Family* 63 (1), 1-16.
- Bengtson, V.L. and Kuypers, J.A. (1971): Generational Difference and Developmental Stake. *Aging and Human Development* 2 (4), 249-260.
- Bengtson, V.L. and Roberts, R.E.L. (1991): Intergenerational Solidarity in Aging Families: An Example of Formal Theory Construction. *Journal of Marriage and the Family* 53 (4), 856-870.
- Bengtson, V.L./Rosenthal, C. and Burton, L. (1990): Families and Aging: Diversity and Heterogeneity. In: Binstock, R.H. and George, L.K. (Eds.): *Handbook of Aging and the Social Sciences*. 3<sup>rd</sup> edition. San Diego.
- Bengtson, V.L./Giarrusso, R./Mabry, J.B. and Silverstein, M. (2002): Solidarity, Conflict, and Ambivalence: Complementary or Competing Perspectives on Intergenerational Relationships? *Journal of Marriage and the Family* 64 (3), 568-576.
- Bertram, H. (2003): Die multilokale Mehrgenerationenfamilie. Von der neolokalen Gattenfamilie zur multilokalen Mehrgenerationenfamilie. In: Feldhaus, M./Logemann, N. and Schlegel, M. (Eds.): *Blickrichtung Familie – Vielfalt eines Forschungsgegenstandes*. Würzburg.

- Bien, W. and Marbach, J.H. (Eds.) (2008): *Familiale Beziehungen, Familienalltag und soziale Netzwerke. Ergebnisse der drei Wellen des Familiensurvey*. Wiesbaden.
- Bösch-Supan, A. and Jürges, H. (Eds.) (2005): *The Survey of Health, Aging and Retirement in Europe (SHARE). Methodology*. Mannheim.
- Bösch-Supan, A./Brugiavini, A./Jürges, H./Mackenbach, J./Siegrist, J. and Weber, G. (Eds.) (2005): *Health, Aging and Retirement in Europe. First Results from the Survey of Health, Aging and Retirement in Europe (SHARE)*. Mannheim.
- Brandt, M. and Szydlik, M. (2008): Soziale Dienste und Hilfe zwischen Generationen in Europa. *Zeitschrift für Soziologie* 37 (4), 301-320.
- Cicirelli, V.G. (1993): Attachment and Obligation as Daughters' Motives for Care-giving Behavior and Subsequent Effect on Subjective Burden. *Psychology and Aging* 8 (2), 144-155.
- Dallinger, U. (2002): Das 'Problem der Generationen': Theorieentwicklung zu inter-generationellen Beziehungen. In: Dallinger, U. and Schroeter, K.R. (Eds.): *Theoretische Beiträge zur Alterssoziologie*. Opladen.
- Dietzel-Papakyriakou, M. (1993): Altern in der Migration. Die Arbeitsmigranten vor dem Dilemma: zurückkehren oder bleiben? Stuttgart.
- Dykstra, P.A. (1999): *Netherlands Kinship Panel Study. A Multi-Actor, Multi-Method Panel Survey on Solidarity in Family Relationships*. The Hague.
- Dykstra, P.A./Kalmijn, M./Knijn, T.C.M./Komter, A.E./Liefbroer, A.C. and Mulder, C.H. (Eds.) (2006): *Family Solidarity in the Netherlands*. Amsterdam.
- Ette, A./Hullen, G./Leven, I. and Ruckdeschel, K. (2007): *Generations and Gender Survey. Dokumentation der Befragung von türkischen Migranten in Deutschland. Materialien zur Bevölkerungswissenschaft, Heft 121b*.
- Frick, J.R. (2006): A General Introduction to the German Socio-Economic Panel Study (SOEP). [http://www.diw.de/documents/dokumentenarchiv/17/43529/soep\\_overview.pdf](http://www.diw.de/documents/dokumentenarchiv/17/43529/soep_overview.pdf). [Last visited 01/18/2010].
- Friedman, D./Hechter, M. and Kreager, D. (2008): A Theory of the Value of Grandchildren. *Rationality and Society* 20 (1), 31-63.
- Giarrusso, R. and Zucker, D. (2004): *The USC Longitudinal Study of Generations. Time-1 Thru Time-7 Codebook*. Ann Arbor.
- Giarrusso, R./Stallings, M. and Bengtson, V.L. (1995): The "Intergenerational Stake" Hypothesis Revisited: Parent-Child Differences in Perception of Relationships 20 Years Later. In: Bengtson, V.L./Schaie, K.W. and Burton, L.M. (Eds.): *Adult Intergenerational Relations. Effects of Societal Change*. New York.
- Giarrusso, R./Feng, D. and Bengtson, V.L. (2004): The Intergenerational-Stake Phenomenon Over 20 Years. *Annual Review of Gerontology and Geriatrics* 24, 55-76.
- Giarrusso, R./Silverstein, M./Gans, D. and Bengtson, V.L. (2005): Aging Parents and Adult Children: New Perspectives on Intergenerational Relationships. In: Johnson, M.L./Bengtson, V.L./Coleman, P.G./Kirkwood, Th.B.L. (Eds.): *Cambridge Handbook of Age and Aging*. London.
- Grünendahl, M. and Martin, M. (2005): Intergenerative Solidarität und praktische Implikationen. In: Otto, U. and Bauer, P. (Eds.): *Mit Netzwerken professionell zusammenarbeiten. Band 1: Soziale Netzwerke in Lebenslauf- und Lebenslagenperspektiven*. Tübingen.

- Haberkern, K. and Szydlik, M. (2008): Pflege der Eltern – Ein europäischer Vergleich. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 60 (1), 78-101.
- Hank, K. (2007): Proximity and Contacts Between Older Parents and Their Children: A European Comparison. *Journal of Marriage and the Family* 69 (1), 157-173.
- Hank, K. and Buber, I. (2009): Grandparents Caring for Their Grandchildren. Findings From the 2004 Survey of Health, Aging, and Retirement in Europe. *Journal of Family Issues* Vol. 30 (1), 53-73.
- Hammarström, G. (2005): The Construct of Intergenerational Solidarity in a Lineage Perspective: A Discussion on Underlying Theoretical Assumptions. *Journal of Aging Studies* 19, 33-51.
- Harper, S. (2005): Grandparenthood. In: Johnson, M.L. (Ed.): *The Cambridge Handbook of Age and Aging*. London.
- Hermalin, A. (2002): *The Well-Being of the Elderly in Asia. A Four-Country Comparative Study*. Ann Arbor.
- Hoff, A. (2006): Intergenerationale Familienbeziehungen im Wandel. In: Tesch-Römer, C./Engstler, H. and Wurm, S. (Eds.): *Altwerden in Deutschland. Sozialer Wandel und individuelle Entwicklung in der zweiten Lebenshälfte*. Wiesbaden.
- Hoff, A. (2007): Patterns of Intergenerational Support in Grandparent-Grandchild and Parent-Child Relationships in Germany. *Aging and Society* 27 (5), 643-665.
- Huinink, J. and Feldhaus, M. (2008): Beziehungs- und Familienentwicklung – eine konzeptionelle Einführung in ein Forschungsprogramm. In: Feldhaus, M. and Huinink, J. (Eds.): *Neuere Entwicklungen in der Beziehungs- und Familienforschung. Vorstudien zum Beziehungs- und Familienentwicklungs-panel (PAIRFAM)*. Würzburg.
- Kalmijn, M. (2008): The Effects of Separation and Divorce on Parent-Child Relationships in Ten European Countries. In: Saraceno, Ch. (Ed.): *Families, Aging and Social Policy. Intergenerational Solidarity in European Welfare States*. Cheltenham/Northampton.
- Katz, R./Lowenstein, A./Phillips, J. and Daatland, S.O. (2005): Theorizing Intergenerational Family Relations. Solidarity, Conflict, and Ambivalence in Cross-National Contexts. In: Bengtson, V.L./Acock, A.C./Allen, K.R./Dilworth-Anderson, P. and Klein, D.M. (Eds.): *Sourcebook of Family Theory and Research*. Thousand Oaks.
- Kaufman, G. and Uhlenberg, P. (1998): Effects of Life Course Transitions on the Quality of Relationships Between Adult Children and Their Parents. *Journal of Marriage and the Family* 60, 924-938.
- King, V. and Elder, G.H. (1995): American Children View Their Grandparents: Linked Lives Across Three Rural Generations. *Journal of Marriage and the Family* 57 (1), 165 – 178.
- King, V. and Elder, G.H. (1997): The Legacy of Grandparenting: Childhood Experiences with Grandparents and Current Involvement with Grandchildren. *Journal of Marriage and the Family* 59 (4), 848-859.
- Klaus, D. (2009): Why Do Adult Children Support Their Parents? *Journal of Comparative Family Studies* 40 (2), 227-241.
- Kohli, M. (2004): Intergenerational Transfers and Inheritance: A Comparative View. In: Silverstein, M. and Schaie, K.W. (Eds.): *Intergenerational Relations Across Time and Place. Annual Review of Gerontology and Geriatrics* 24. New York.

- Kohli, M. and Künemund, H. (2005a): The Midlife Generation in the Family: Patterns of Exchange and Support. In: Willis, S.L. and Martin, M. (Eds.): *Middle Adulthood: A Lifespan Perspective*. Thousand Oaks/London/New Delhi.
- Kohli, M. and Künemund, H. (Eds.) (2005b): *Die zweite Lebenshälfte. Gesellschaftliche Lage und Partizipation im Spiegel des Alters-Survey*. 2<sup>nd</sup> edition. Wiesbaden.
- Kohli, M./Künemund, H. and Lüdicke, J. (2005): Family Structure, Proximity and Contact. In: Bösch-Supan, A./Brugiavini, A./Jürges, H./Mackenbach, J./Siegrist, J. and Weber, G. (Eds.): *Health, Aging and Retirement in Europe. First Results from the Survey of Health, Aging and Retirement in Europe (SHARE)*. Mannheim.
- Komter, A. and Schans, D. (2008): Reciprocity Revisited: Give and Take in Dutch and Immigrant Families. *Journal of Comparative Family Studies* 39 (2), 279-298.
- Künemund, H. (2006): Changing Welfare States and the „Sandwich Generation“. Increasing Burden for the Next Generation? *International Journal of Aging and Later Life* 1 (2), 11-29.
- Künemund, H. (2008): Intergenerational Relations Within the Family and the State. In: Saraceno, Ch. (Ed.): *Families, Aging and Social Policy. Intergenerational Solidarity in European Welfare States*. Cheltenham/Northampton.
- Künemund, H. and Vogel, C. (2006): Öffentliche und private Transfers und Unterstützungsleistungen im Alter – „crowding out“ oder „crowding in“? *Zeitschrift für Familienforschung* 18 (3), 269-289.
- Lauterbach, W. (1995): Die gemeinsame Lebenszeit von Familiengenerationen. *Zeitschrift für Soziologie* 24 (1), 22-41.
- Lauterbach, W. (2004): *Die multilokale Mehrgenerationenfamilie. Zum Wandel der Familienstruktur in der zweiten Lebenshälfte*. Würzburg.
- Lauterbach, W. and Klein, Th. (2004): The change of generational relations based on demographic developments: The case of Germany. *Journal of Comparative Family Studies* 35 (4), 651-663.
- Lauterbach, W. and Lüscher, K. (1996): Erben und die Verbundenheit der Lebensverläufe von Familienmitgliedern. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 48 (1), 66-95.
- Lawton, L./Silverstein, M. and Bengtson, V.L. (1994): Affection, Social Contact and Geographic Distance between Adult Children and Their Parents. *Journal of Marriage and the Family* 56 (1), 57-68.
- Lin, I-F. (2008): Consequences of Parental Divorce for Adult Children's Support of Their Frail Parents. *Journal of Marriage and the Family* 70 (1), 113-128.
- Lowenstein, A. and Ogg, J. (2003): *OASIS – Old Age and Autonomy: The Role of Service Systems and Intergenerational Family Solidarity. Final Report*. Haifa.
- Lüscher, K. and Pillemer, K. (1998): *Intergenerational Ambivalence: A New Approach to the Study of Parent-Child Relations in Later Life*. *Journal of Marriage and the Family* 60 (2), 413-425.
- Mangen, D.J./Bengtson, V.L. and Landry, P.H. (Eds.) (1988): *The Measurement of Intergenerational Relations*. Newbury Park/Beverly Hills/London/New Delhi.
- Merz, E.-M./Schuengel, C. and Schulze, H.-J. (2007): Intergenerational Solidarity: An Attachment Perspective. *Journal of Aging Studies* 21, 175-186.

- Merz, E.-M./Schuengel, C. and Schulze, H.-J. (2008): Inter-generational Relationships at Different Ages: An Attachment Perspective. *Aging and Society* 28, 717-736.
- Mueller, M.M. and Elder, G.H. (2003): Family Contingencies Across the Generations: Grandparent-Grandchild Relationships in Holistic Perspective. *Journal of Marriage and Family* 65 (2), 404-417.
- Nauck, B. (2001): Social Capital, Intergenerational Transmission and Intercultural Contact in Immigrant Families. *Journal of Comparative Family Studies* 32, 465-488.
- Nauck, B. (2007): Immigrant Families in Germany. Family change between situational adaptation, acculturation, segregation and remigration. *Zeitschrift für Familienforschung* 19 (1), 34-54.
- Nauck, B. (2009a): Patterns of Exchange in Kinship Systems in Germany, Russia, and the People's Republic of China. *Journal of Comparative Family Studies* 40 (2), 255-278.
- Nauck, B. (2009b): Intergenerational Relationships and Female Inheritance Expectations. Comparative Results from Seven Societies in Asia and Europe. *Asian Journal of Social Psychology* (12).
- Nauck, B. and Kohlmann, A. (1998): Verwandtschaft als soziales Kapital – Netzwerkbeziehungen in türkischen Migrantenfamilien. In: Wagner, M. and Schütze, Y. (Eds.): *Verwandtschaft. Sozialwissenschaftliche Beiträge zu einem vernachlässigten Thema*. Stuttgart.
- Nauck, B. and Suckow, J. (2006): Intergenerational Relationships in Cross-Cultural Comparison: How Social Networks Frame Intergenerational Relations Between Mothers and Grandmothers in Japan, Korea, China, Indonesia, Israel, Germany, and Turkey. *Journal of Family Issues* 27 (8), 1159-1185.
- Nauck, B. and Yi, Ch.-Ch. (Eds.) (2007): Intergenerational Relationships in Cross-Cultural Perspective: Fertility, Interaction and Support. Special Issue of *Current Sociology* 55.
- Parrott, T.M. and Bengtson, V.L. (1999): The Effects of Earlier Intergenerational Affection, Normative Expectations, and Family Conflict on Contemporary Exchange of Help and Support. *Research on Aging* 21 (1), 73-105.
- Pillemer, K. and Lüscher, K. (Eds.) (2004): *Intergenerational Ambivalences. New Perspectives on Parent-Child Relations in Later Life*. Oxford.
- Pillemer, K. and Suitor, J.J. (2002): Explaining Mothers' Ambivalence Toward Their Adult Children. *Journal of Marriage and the Family* 64 (3), 602-613.
- Roberts, R.E.L. and Bengtson, V.L. (1990): Is Intergenerational Solidarity a Unidimensional Construct? A Second Test of a Formal Model. *Journal of Gerontology* 45 (1), 12-20.
- Rossi, A.S. (1993): Intergenerational Relations: Gender, Norms, and Behavior. In: Bengtson, V.L. and Achenbaum, W.A. (Eds.): *The Changing Contract Across Generations*. New York.
- Rossi, A.S. and Rossi, P.H. (1990): *Of Human Bounding. Parent-Child Relations Across the Life Course*. New York.
- Ruckdeschel, K./Ette, A./Hullen G. and Leven, I. (2006): Generations and Gender Survey. Dokumentation der ersten Welle der Hauptbefragung in Deutschland. Materialien zur Bevölkerungswissenschaft, Heft 121a.



- Schwarz, B. and Trommsdorff, G. (2005): The Relation Between Attachment and Intergenerational Support. *European Journal of Aging* 2 (3), 192-199.
- Silverstein, M./Lawton, L. and Bengtson, V.L. (1994): Types of Relations Between Parents and Adult Children. In: Bengtson, V.L. and Harootyan, R.A. (Eds.): *Intergenerational Linkages. Hidden Connections in American Society*. New York.
- Silverstein, M./Conroy, St.J./Wang, H./Giarrusso, R. and Bengtson, V.L. (2002): Reciprocity in Parent-Child Relations Over the Adult Life Course. *Journal of Gerontology: Social Sciences*: 57B, 3-11.
- Spitze, G. and Logan, J.R. (1991): Sibling Structure and Intergenerational Relations. *Journal of Marriage and the Family* 53, 871-884.
- Steinbach, A. (2008): Intergenerational Solidarity and Ambivalence: Types of Relationships in German Families. *Journal of Comparative Family Studies* 39 (1), 115-127.
- Steinbach, A. and Kopp, J. (2008a): Intergenerationale Beziehungen. Theoretische Diskussionen, empirische Befunde und offene Fragen. In: Feldhaus, M. and Huinink, J. (Eds.): *Neuere Entwicklungen in der Beziehungs- und Familienforschung. Vorstudien zum Beziehungs- und Familienentwicklungspanel (PAIRFAM)*. Würzburg.
- Steinbach, A. and Kopp, J. (2008b): "When will I see you again?" Intergenerational Contacts in Germany. In: Saraceno, Ch. (Ed.): *Families, Aging and Social Policy. Intergenerational Solidarity in European Welfare States*. Cheltenham/Northampton.
- Sweet, J.A. and Bumpass, L.L. (2002): *The National Survey of Families and Households – Waves 1, 2, and 3: Data Description and Documentation*. Center for Demography and Ecology, University of Wisconsin-Madison. <http://www.ssc.wisc.edu/nsfh/home.htm>. [Last visited 01/18/2010].
- Szydlik, M. (1995): Die Enge der Beziehung zwischen erwachsenen Kindern und ihren Eltern – und umgekehrt. *Zeitschrift für Soziologie* 24 (2), 75-94.
- Szydlik, M. (1999): Erben in der Bundesrepublik Deutschland. Zum Verhältnis von familialer Solidarität und sozialer Ungleichheit. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 51 (1), 80-104.
- Szydlik, M. (2000): *Lebenslange Solidarität? Generationenbeziehungen zwischen erwachsenen Kindern und Eltern*. Opladen.
- Szydlik, M. (2004): Inheritance and Inequality: Theoretical Reasoning and Empirical Evidence. *European Sociological Review* 20 (1), 31-45.
- Szydlik, M. (2008): Intergenerational Solidarity and Conflict. *Journal of Comparative Family Studies* 34 (1), 97-114.
- Szydlik, M. and Schupp, J. (2004): Wer erbt mehr? Erbschaften, Sozialstruktur und Alterssicherung. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 56 (4), 609-629.
- Tesch-Römer, C./Engstler, H. and Wurm, S. (Eds.) (2006): *Altwerden in Deutschland. Sozialer Wandel und individuelle Entwicklung in der zweiten Lebenshälfte*. Wiesbaden.
- Tesch-Römer, C./Kondratowitz, H.-J. von/Motel-Klingebiel, A. and Spangler, D. (2000): *OASIS – Old Age and Autonomy: The Role of Service Systems and*

- Intergenerational Family Solidarity. Erhebungsdesign und Instrumente des deutschen Surveys. Berlin.
- Tesch-Römer, C./Wurm, S./Hoff, A. and Engstler, H. (2002): Die zweite Welle des Alterssurveys. Erhebungsdesign und Instrumente. DZA Diskussionspapiere, Nr. 35.
- Trommsdorff, G. and Nauck, B. (Eds.) (2005): The Value of Children in Cross-cultural Perspective. Lengerich.
- Trommsdorff, G. and Schwarz, B. (2007): The „Intergenerational Stake Hypothesis“ in Indonesia and Germany: Adult Daughters' and their Mothers' Perception of their Relationship. *Current Sociology* 55, 599-620.
- Van Gaalen, R.I. and Dykstra, P.A. (2006): Solidarity and Conflict Between Adult Children and Parents: A Latent Class Analysis. *Journal of Marriage and the Family* 68 (4), 947-960.

## 5.5 Administrative Data from Germany's Statutory Health Insurance Providers for Social, Economic, and Medical Research

*Günther Heller*

Contact:

Günther Heller  
Research Institute of the Local Healthcare Funds (WIdO)  
Rosenthaler Straße 31  
10178 Berlin  
Germany  
e-mail: [guenther.heller\[at\]wido.bv.aok.de](mailto:guenther.heller[at]wido.bv.aok.de)

## **Abstract**

This article gives a short description of the administrative data sources used by Germany's statutory health insurance providers. These data sources are of potential interest for social, economic, and medical research. We first briefly outline the legal regulations applying to these sources, the structure and contents of the most relevant databases, as well as current and future access to these data sources in the context of legal and data privacy protection issues. We then discuss issues of data validity and completeness of different data sources in relation to a recent example of health care research using administrative data. In conclusion, we discuss the potential and limitations of research using administrative medical data from Germany's statutory health insurance providers.

Keywords: administrative data, claims data, health service research

## **1. Introduction**

Recent years have seen growing importance of administrative medical data for scientific purposes. Clinical research, health service research, and the health care economy all increasingly demand medical data not only from surveys and clinical trials, but also data on entire populations. This demand stems from a need to analyze and evaluate medical innovation and developments more broadly based upon real-life settings. In light of this, it is only fitting to address the topic of administrative medical data in terms of its structure, availability, and usefulness.

To avoid confusion about the term "administrative medical data," the following short description is given: the term refers to data that has been established for administrative purposes in the field of health care provision. In most cases, administrative medical data has been collected for the reimbursement of health care providers (e.g., doctors, nurses, or hospitals), for usage in official statistics, and other administrative purposes. Thus, in the context of scientific research, administrative medical data are secondary data. Other frequently used or similar terms are "administrative data," "claims data," or "reimbursement data." Administrative medical data may be contrasted with data from medical surveys, medical trials, or data derived directly from medical records.

There are numerous examples of administrative medical data from the international field. Some well-known examples are the Hospital Episodes

Statistics (HES) from Great Britain,<sup>1</sup> the Veterans Affairs data collections, or the Medicare databases from the US.<sup>2</sup>

In this advisory report, the focus will be on the administrative medical data of statutory health insurance providers in Germany. The structure and availability of these data depend on which part of the health care sectors they pertain to. An example from the field of health service research will be provided and future developments will be discussed.

## **2. Administrative medical data of German statutory health insurances**

For 125 years, medical care in Germany has been financed through the health insurance system. Currently, approximately 86 percent of the population is insured under the statutory health insurance system (GKV, *Gesetzliche Krankenversicherung*), while the remaining 14 percent is insured by private health insurance (PKV, *Private Krankenversicherung*) (Jacobs et al. 2006). Health insurance is largely governed by the Code of Social Law V (SGB V), however several smaller parts are covered by the Code of Social Law IX (Rehabilitation) (SGB IX) and XI (Nursing) (SGB XI). The collection of routine medical data, as well as its transfer to statutory health insurance providers,<sup>3</sup> is also regulated by these Codes.

The provision of health care in Germany is divided into different sectors, each of which is regulated individually. For the sectors listed in table 1, health care providers are obliged to provide individual-specific medical data when submitting an invoice for reimbursement from the statutory health insurance. It is, however, unclear whether all statutory health insurance providers have established databases that allow these data from different health sectors to be linked at an (anonymous) individualized level. Table 1 provides an overview of the different laws pertaining to different health care sectors according to SGB V, as well as examples of the most important medical data required to be transferred to statutory health insurance providers. Additionally, a list of the currently available anonymous individualized-level databases within the Research Institute of the Local Healthcare Funds (WiDo, *Wissenschaftliches Institut der AOK*) is provided. The WiDo is responsible for the databases for German regional healthcare funds (AOK, *Allgemeine Ortskrankenkasse*). Currently, approximately 24 million people

---

1 <http://www.hesonline.nhs.uk/Ease/servlet/ContentServer?siteID=1937&categoryID=537>, cited on 09/27/2008.

2 An overview of “coded data from administrative sources” in the US is provided by Iezzoni (2003).

3 Or maybe different insurance providers when referring to SGB IX and XI.

are insured by the AOK, making it – compared with other health insurance providers – the largest database of this kind in Germany.

Table 1: Health Care Provision according to Code of Social Law V (SGB V) in different medical sectors and the data transfer to statutory health insurance providers in Germany

	Medical Sector	Main Medical Contents (Examples)	WIdO Database since (2)
<b>SGB V § 295</b>	Ambulatory care	Diagnosis, type of medical care provided	2004
<b>SGB V § 295</b>	Incapable of work	Diagnosis, time being incapable of working	2006
<b>SGB V § 300</b>	Pharmaceutical prescription	type of pharmaceutical product, price, quantity of prescribed pharmaceutical	1998
<b>SGB V § 301</b>	Hospital care	Admitted hospital, diagnoses, operations, procedures, length of stay	1998
<b>SGB V § 302</b>	Prescription of remedies and medical aids (1)	Diagnosis, type, and quantity of medical care provided, price	2004
<b>SGB XI § 93 – 98</b>	Nursing Care	Diagnoses, type of care	-
<b>SGB IX</b>	Rehabilitation	Admitted hospital, diagnoses, operations, procedures, length of stay	2004

(1) e.g., massages, ergo therapy, physical therapy, prosthesis, etc.

(2) case specific, able to be linked on a anonymous individual level

In general, the described administrative medical data contain the following information, which are usually stored in Structured Query Language (SQL) databases:

- personal identifier
- date(s) of medical care provision (episode)
- type of disease (e.g., International Statistical Classification of Diseases and Related Health Problems, also known as ICD)
- type of treatment (e.g., procedural classification)
- invoice
- other information

These data may be linked to additional administrative data related to the insured individual. Therefore, information about, for example, the individual's place of residence, status of insurance, or end of insurance may be used for further scientific purposes.

In light of this, information collected from medical sectors and during an individual's contact with the medical system has potential to be used for individual longitudinal analyses. This would be of great interest for research purposes.

## *2.1 Accessibility of data*

It may be seen as a serious drawback that the data described is currently only available to health insurance providers and to researchers performing research in cooperation with these health insurance providers (AOK-Bundesverband et al. 2007; Bramesfeld et al. 2007; Geyer 2008; Grobe et al. 2008; Heller et al. 2004; Heller 2006; Heller 2007; Heller et al. 2007; Ihle et al. 2005; Müller and Braun 2006; Schubert et al. 2007; Swart and Heller 2007; Swart and Ihle 2005). This limited data access is a direct result of the protection of the private data of insured individuals and the involved institutions, such as hospitals. However, there are currently at least two regulations which have the potential to make these data available to the scientific community in the future; namely, § 303 SGB V Data Transparency and § 137 SGB V Quality of Medical Care.

According to § 303 SGB V Data Transparency, administrative medical data from all health insurance providers is to be pooled in a data trust center. These data should be made available to health insurers, health authorities, and several other defined user groups, such as, for example, independent scientific organizations. However, the level of aggregation at which the data can be made available is not fixed. Thus, data may be distributed on an individual level, or on an aggregated level. However, it remains unclear how expensive it would be to use these data for external research. How the data are to be compiled and distributed, and at which expense they are to be provided, are issues still being debated.

In the latest health care reform of March 2007, § 137 SGB V Quality of Medical Care was renewed. According to this reform, a quality agency shall be assigned to develop and provide “inter-sectoral” comprehensive quality assurance. The agency that is to be developed for this purpose has additionally been granted permission to use administrative medical data. Consequently, for future considerations relating to administrative medical data, it may be of interest to direct attention towards this agency. Which institution will assume the role of this quality agency is, however, still unclear, as the pan-European tendering process for this agency is still in progress.

## *2.2 Aspects of data quality*

For numerous research-related enquiries, one central question relating to administrative medical data is whether the required information is present in the data. Interesting information might or might not be sufficiently present. For example, the social status of an individual or issues pertaining to an individual’s quality of life often play an important role in a person’s health, however, administrative medical data often do not include such information.



Currently, for example, there is much debate as to whether administrative medical data provides sufficient information to perform risk-adjusted analyses when comparing hospital performance measurements, such as, for example, 30-day survival rates after diagnosis of acute myocardial infarction. While the effectiveness of risk-adjusted analyses using administrative medical data has been questioned for some time, recent research from Great Britain and the US reveals similar performance of administrative medical data compared to clinical data or clinical register data when predicting survival after admission to hospital due to tracer-diagnoses or procedures (Aylin et al. 2007; Pine et al. 2007). It should, however, be noted that good prediction is only a poor indicator for good risk adjustment (Heller and Schnell 2007).

Another important issue related to administrative medical data is data validity. It should be noted that while administrative medical data is secondary data, how valid the data is depends on the primary purpose for which the data was collected. For example, reimbursement data from the hospital sector is usually checked by health insurance providers in terms of the accuracy of invoices. This is done using plausibility checks (internal validity). In addition, health insurance providers also conduct audits comparing the transferred administrative medical data with clinical data from medical records (external validity). When scrutinizing reimbursement data, several data elements are typically considered unimportant. Items such as time of coded procedures or admission diagnoses are, for example, neither checked nor corrected. Nevertheless, such information may be of great importance from various analytical perspectives. Thus, when performing medical administrative data analyses for various purposes, one should take note of the data available as well as its validity.

To my knowledge, there have hitherto been no studies from any German health care sector examining the validity of administrative medical data from a medical or scientific perspective. Before using any administrative medical data for such purposes, first undertaking an external study would seem appropriate. Generally, it is reasonable to assume that the validity of administrative medical data increases with time as administrative procedures regarding data transfer become more established. In light of this, administrative medical data from hospital care or pharmaceutical prescriptions – in both of which areas data transfer practices were established in 1998 – are usually considered valid and reliable. Data transfers from the ambulatory sector, on the other hand, were not established until 2004 and have thus been called into question (Gerste and Gutschmidt 2006; Giersiepen et al. 2007; Trautner et al. 2005).

An interesting matter connected to data validity is whether information that may be of potential interest to researchers is linked to reimbursement. For example, the type and amount of prescribed pharmaceuticals is directly

linked to reimbursement, which is thoroughly examined by insurers and thus deemed to be of a high level of validity. Documentation of primary and secondary diagnoses in the hospital sector, which are used to verify data validity, has dramatically increased since the implementation of Diagnosis Related Group (DRG) hospital reimbursement schemes. Diagnoses from the ambulatory sector, on the other hand, are not directly related to reimbursement, not scrutinized by insurers, and have accordingly been shown to be invalid in many cases (Gerste and Gutschmidt 2006; Giersiepen et al. 2007; Trautner et al. 2005). However, in light of the recent “morbidity-orientated risk-structure compensation” (*Morbiditäts-Orientierter-Risikostrukturausgleich*), this may change considerably in the near future.

In addition to the validity of the data, one must also take the completeness of the data into consideration. Administrative medical data is typically considered to be complete when all relevant cases are present in the data. Administrative medical data need to be comprehensive, as these data are utilized by several different health care sector bodies for various purposes, such as, for example, the DRG-based statistics on hospital diagnoses compiled by the Federal Statistical Office (Spindler 2008). Another example is the use of administrative medical data by the Federal Office for Quality Assurance (BQS, *Institut für Qualität und Patientensicherheit*) to create quality benchmarks (Veith et al. 2008).

Several years ago, a working group entitled the Working Group for Secondary Data Analysis (AGENS, *Arbeitsgruppe Erhebung und Nutzung von Sekundärdaten*) was created to assess the contents and possible uses of administrative medical data. In addition to AGENS, a handbook was also created to provide a detailed overview of administrative medical data in Germany (Swart and Ihle 2005). A more compact, updated version of this handbook is currently available (Swart and Heller 2007). In addition to this handbook, the same authors have developed a set of guidelines addressing “Good Practice for Secondary Data Analysis” (Swart et al. 2005). Since these guidelines were formulated, they have been adopted by several scientific societies as well as funding agencies.<sup>4</sup>

---

4 [http://www.gesundheitsforschung-bmbf.de/\\_media/GPS.pdf](http://www.gesundheitsforschung-bmbf.de/_media/GPS.pdf), cited 02/10/2008.

### **3. A health service research example using administrative medical data**

As mentioned, administrative medical data can be used for various research purposes. One such example is the “volume outcome relationship” of very low birth weight infants (VLBW). In Germany, the issue of whether a minimum level of provision be introduced for hospitals treating VLBW infants has been addressed by a Federal Joint Committee (Gemeinsamer Bundesausschuss). Before the decision of the Federal Joint Committee was reached, it commissioned the Institute for Quality and Efficiency in Health Care (IQWiG, *Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen*) to evaluate whether a “volume outcome relationship” for this patient group exists. In its final report, the IQWiG used analyses by Heller (2007), based upon administrative Hospital claims data (IQWiG 2008). These analyses have since been extended to include spatial simulation analyses for different “minimum provider volumes.” The analyses conducted by Heller first measured distances from patients’ residences to hospitals providing care. Second, a simulation was conducted measuring the extent to which these distances changed after “minimum provider volumes” were introduced. Additionally, the changes in survival rates after introducing various levels of “minimum provider volumes” were estimated (Heller 2009). These and several other similar analyses are currently being used by the Federal Joint Committee.

### **4. Discussion and conclusion**

- Three points on administrative medical data in Germany are to be noted:
- administrative medical data provide detailed information about medical care provision for large samples or even entire populations;
- these data provide information about existing diseases, medical therapy, and health outcomes; and
- researchers with clinical economic and social research interests have shown particular interest in performing individual longitudinal analyses using individual identifiers.

With this article, I aimed to give the reader a brief description of the most important administrative medical databases in Germany, and to address issues like data validity, completeness, and accessibility of administrative

medical data. Additionally, to illustrate the importance of administrative medical data, I provided an up-to-date example of its use in health service research on very low birth weight infants.

One point which should be kept in mind, however, is that health and illness also exists outside the official medical system within the lay system (Borgetto and Trojan 2007). The onset of a disease might occur long before a doctor is contacted or self-administered treatments are undertaken (such as using readily available over-the-counter drugs, which have been known to cure numerous diseases without the patient ever coming into contact with the official medical system). Incidence or prevalence studies may be difficult to justify under this restriction.

From this perspective, the administrative medical data of statutory health insurance providers tell only part of the story. Thus, depending on the research question, it is in most cases desirable to complement administrative medical data with other data sources.

## References:

- AOK-Bundesverband, Forschungs- und Entwicklungsinstitut für das Sozial- und Gesundheitswesen in Sachsen Anhalt (FEISA), HELIOS Kliniken, Wissenschaftliches Institut der AOK (WiDO), Editor. Qualitätssicherung der stationären Versorgung mit Routinedaten (QSR).
- Aylin, P./Bottle, A. and Majeed, A. (2007): Use of administrative data or clinical databases as predictors of risk of death in hospital: comparison of models. *BMJ* 334 (7602), 1044.
- Borgetto, B./Trojan, A. (2007): Versorgungsforschung und Laiensystem. In: Janßen, C./Borgetto, B. and Heller, G. (Eds.): Medizinsoziologische Versorgungsforschung. Theoretische Ansätze, Methoden, Instrumente und empirische Befunde. Weinheim.
- Bramesfeld, A./Grobe, T. and Schwartz, F.W. (2007): Who is treated, and how, for depression? An analysis of statutory health insurance data in Germany. *Soc Psychiatry Psychiatr Epidemiol* 42 (9), 740-746.
- Gerste, B. and Gutschmidt, S. (2006): Datenqualität von Diagnosen aus dem ambulanten Bereich. *Gesundheits- und Sozialpolitik* 2006, (3-4), 29-43.
- Geyer, S. (2008): Social inequalities in the incidence and case fatality of cancers of the lung, the stomach, the bowels, and the breast. *Cancer Causes Control* 19 (9), 965-974.
- Giersiepen, K./Pohlabeln, H./Egidi, G. and Pigeot, I. (2007): Die ICD-Kodierqualität für Diagnosen in der ambulanten Versorgung. *Bundesgesundheitsblatt – Gesundheitsforschung – Gesundheitsschutz* 8, 1028-1038.
- Grobe, T.G./Gerhardus, A./A'Walelu, O./Meisinger, C. and Krauth, C. (2008): [Hospitalisations for acute myocardial infarction – comparing data from three different sources]. *Gesundheitswesen* 70 (8-9), e37-46.
- Heller, G./Swart, E. and Mansky, T. (2004): Qualitätsanalysen mit Routinedaten. Ansatz und erste Analysen aus dem Gemeinschaftsprojekt "Qualitätssicherung mit Routinedaten" (QSR). In: Klauber, J./Robra, B.P. and Schellschmidt, H. (Eds.): Krankenhaus-Report 2003. Stuttgart/New York.
- Heller, G. (2006): Sind risikoadjustierte Analysen mit administrativen Routinedaten möglich? In: Hey, M. and Maschewski-Schneider, U. (Eds.): Kursbuch Versorgungsforschung. Berlin.
- Heller, G./Günster, C./Misselwitz, B./Feller, A. and Schmidt, S. (2007): Jährliche Fallzahl pro Klinik und Überlebensrate sehr untergewichtiger Frühgeborener (VLBW) in Deutschland. Eine bundesweite Analyse mit Routinedaten. *Z Geburtshilfe Neonatol* 211 (3), 123-131.
- Heller, G. and Schnell, R. (2007): Hospital Mortality Risk Adjustment Using Claims Data. *JAMA* 297 (18), 1983.
- Heller, G. (2009): Auswirkungen der Einführung von Mindestmengen in der Behandlung von sehr untergewichtigen Früh- und Neugeborenen (VLBW). Eine Simulation mit Echtdaten. In: Klauber, J./Robra, B. and Schellschmidt, H. (Eds.): Krankenhaus-Report 2008. Stuttgart.
- Ihle, P./Köster, I./Herholz, H./Rambow-Bertram, P./Schardt, T. and Schubert, I. (2005): [Sample survey of persons insured in statutory health insurance

- institutions in Hessen – concept and realisation of person-related data base]. *Gesundheitswesen* 67 (8-9), 638-645.
- Iezzoni, L.I. (2003): Coded data from administrative sources. In: Iezzoni, L.I. (Ed.): *Risk adjustment for measuring health care outcomes*. 3rd ed. Chicago.
- IQWiG (2008): <http://www.iqwig.de/index.681.html>.
- Jacobs, K./Klauber, J. and Leinert, J. (2006): *Fairer Wettbewerb oder Risikoselektion. Analysen zur gesetzlichen und privaten Krankenversicherung*. Bonn.
- Müller, R. and Braun, B. (Eds.): *Vom Quer zum Längsschnitt. Möglichkeiten der Analysen mit GKV-Daten*. St. Augustin.
- Pine, M./Jordan, H.S./Elixhauser, A./Fry, D.E./Hoaglin, D.C./Jones, B. et al. (2007): Enhancement of Claims Data to Improve Risk Adjustment of Hospital Mortality. *JAMA* 297 (1), 71-76.
- Schubert, I./Kupper-Nybelen, J./Ihle, P. and Krappweis, J. (2007): [Utilization patterns of dementia patients in the light of statutory health insurance data]. *Z Arztl Fortbild Qualitatssich* 101 (1), 7-13.
- Spindler, J. (2008): Fallpauschalenbezogene Krankenhausstatistik: Diagnosen und Prozeduren der Patienten auf Basis der Daten nach § 21 Krankenhausentgeltgesetz. In: Klauber, J./Robra, B. and Schellschmidt, H. (Eds.): *Krankenhaus-Report 2007*. Stuttgart.
- Swart, E. and Heller, G. (2007): Nutzung und Bedeutung von (GKV-)Routinedaten für die Versorgungsforschung. In: Janßen, C./Borgetto, B. and Heller, G. (Eds.): *Medizinsoziologische Versorgungsforschung. Theoretische Ansätze, Methoden, Instrumente und empirische Befunde*. Weinheim.
- Swart, E. and Ihle, P. (Eds.): *Routinedaten im Gesundheitswesen. Handbuch Sekundärdatenanalyse: Grundlagen, Methoden und Perspektiven*. Bern.
- Swart, E./Ihle, P./Geyer, S./Grobe, T. and Hofmann, W. (2005): GPS – Gute Praxis Sekundärdatenanalyse. Arbeitsgruppe Erhebung und Nutzung von Sekundärdaten (AGENS) der Deutschen Gesellschaft für Sozialmedizin und Prävention (DGSMP) [GPS – good practice secondary data analysis. Working Group for the Survey and Utilization of Secondary Data (AGENS) of the German Society for Social Medicine and Prevention (DGSMP)]. *Gesundheitswesen* 67 (6), 416-421.
- Trautner, C./Dong, Y./Ryll, A. and Stillfried, D.G. (2005): Verlässlichkeit von Diagnosen niedergelassener Ärzte in Niedersachsen. *Gesundheits- und Sozialpolitik* (1-2), 36-42.
- Veith, C./Bauer, J./Döbler, K./Eckert, O./Fischer, B. and Woldenga, C. (2008): *Qualität sichtbar machen. BQS-Qualitätsreport 2007*. Düsseldorf.

## 5.6 Provision for Old Age: National and International Survey Data to Support Research and Policy on Aging

*Hendrik Jürges*

Contact:

Hendrik Jürges  
Mannheim Research Institute for the Economics of Aging  
University of Mannheim  
L13, 17  
68131 Mannheim  
Germany  
e-mail: [juerges\[at\]mea.uni-mannheim.de](mailto:juerges[at]mea.uni-mannheim.de)



## **Abstract**

This report reviews recent trends in the collection of multidisciplinary and longitudinal data in the area of aging research, both in Germany and internationally. It also discusses important developments such as linkage with administrative records, the inclusion of health measurements and biomarkers, and the inclusion of populations in institutions, particularly nursing homes.

## **1. Research questions**

Population aging is one of the megatrends of the twenty-first century. In almost all countries of the developed world, mortality rates are falling, birth rates are below replacement rate, and work rates are falling. World wide, the number of people aged 60 and over is expected to triple until 2050. The aging of the population will shape the world to come and its political agenda. The main policy issues that arise with an aging population concern providing income and health security during old age at affordable budgets. To cope with these particular challenges of rapid population aging, it is important to improve the current scientific understanding of complex linkages between economic, health, and social factors that determine the quality of life in the older population. These interactions primarily take place at the individual level, they are dynamic – aging is a process, not a state – and are related to a country's welfare regime.

To improve our understanding of population aging and its policy implications, researchers need multidisciplinary and longitudinal data. Over the past decade, the international research community has responded to this need by starting to create a worldwide microdata infrastructure that helps researchers to better understand the individual and population aging processes. The aim of this special report is to document the contents and degrees of access to existing national and international datasets containing household- and individual-level information on the economic well-being and health of older populations. In this context, the international perspective is of great importance. Internationally comparative data is necessary in order to exploit the rich variety in policies, institutions, and other factors across different countries. The impact of public policy can be much better understood if one can observe one policy in relation to other policies. Many of the policies that might be considered as solutions to address future public policy challenges resulting from an aging population have already been implemented in some form in some country. For instance, comparisons of different pension systems and their impact on old age poverty rates, savings decisions of the

working population, or the role of the family and intergenerational relations, can inform policy-makers about the likely consequences of pension reform.

In addition to briefly describing the contents of and access to existing survey datasets, the present report also discusses future developments and further needs in the area of aging research. Three important areas in which such developments are likely and needed are (1) linkage with administrative records, (2) the inclusion of (more) health measurements and biomarkers, and (3) the inclusion of populations in institutions, particularly nursing homes.

## **2. Status quo: Databases and access**

This section briefly describes a selection of important national and international databases for multidisciplinary research on aging. Summary information and details on data access are given in tabular form in the appendix. Only databases that fulfill several criteria are listed. First, they must of course include – although not exclusively – coverage of the older population (defined as individuals aged 50 and older). Second, they must be based on non-administrative *research-driven* surveys. Third, they must be multidisciplinary; that is, they must include data with some detail from at least two of the following fields: medicine, economics, sociology, and psychology. Fourth, they must have a longitudinal design, because many events associated with aging are dynamic or longitudinal in nature. Obviously, these four criteria together seem to be fairly restrictive. For instance, they exclude the European Community Household Panel (ECHP), many of the health interview surveys carried out on national levels, and also the current European Health Interview Survey (EHIS).

### *2.1 National data sources for Germany*

*German Socio-Economic Panel (SOEP, Sozio-oekonomisches Panel):* arguably, the most important data source of longitudinal microdata is the SOEP, which has collected detailed annual data mainly on housing, employment, and income since 1984. One advantage of the SOEP is the large sample size. The sample currently includes about 22,000 respondents of which some 10,500 are aged 50 and above. Another advantage is the length of the panel. The SOEP now spans 25 years of annual data on the lives of a substantial number of respondents, following individuals from middle age into old age. The usefulness of the SOEP for certain aspects of aging research, however, has been limited by two facts: first, until recently, it

contained little health-related information. For instance, self-rated general health has been included only in 1992, and information on respondents' smoking behavior, height, and weight has been only available since the late 1990s. However, a more systematic approach to the measurement of health has been taken since 2002, when SOEP included a version of the SF-12 health questionnaire. A more recent development has been the inclusion of physical measures or biomarkers, such as handgrip strength (in 2006). Second, although some broad information on savings and household assets is available annually, the quantitative composition of household assets was covered only in 1998, 2002, and 2007, making it difficult to track in detail changes in the asset portfolios or in the amount of wealth.

*The German Ageing Survey (DEAS, Deutscher Alterssurvey)*: DEAS is a cross-sectional and longitudinal survey of individuals aged 40 and over. The main topics it covers address circumstances and attitudes in later life, particularly related to housing, employment and retirement, volunteering, family, leisure activities, and social participation. Data collection started in 1996 and is repeated every six years. In the first two waves, samples included about 5000 respondents. One advantage of the DEAS is that it collects highly detailed information from the respondents, particularly around psychological variables, whereas panel mortality due to the large gap between waves clearly is a disadvantage. Only 32 percent of the original sample members in 1996 could be reinterviewed in 2002. Also, data on the oldest-old (85+) are only available for reinterviewed panel data members, since the baseline sample in 1996 was restricted to individuals up to age 85.

*Sparen und Altersvorsorge in Deutschland (SAVE)*: SAVE is the most elaborate survey on savings behavior in Germany. It contains detailed factual information on the current financial situation of households (including asset components), savings behavior, psychological determinants of saving, and health. SAVE was started in 2001 and data for several waves are now available. One drawback of the data for longitudinal analyses is the somewhat complex mixture of different samples that comprise it, which limits the exploitable longitudinal dimension.

## 2.2 International data sources

The *Cross-National Equivalent File (CNEF)* is an *ex post* harmonized dataset containing annual panel data from the US Panel Study of Income Dynamics (PSID), the German Socio-Economic Panel (SOEP), the British Household Panel Study (BHPS), the Household Income and Labour Dynamics in Australia (HILDA), and the Canadian Survey of Labour and Income Dynamics (SLID). Future releases of the CNEF will include data from the Swiss Household Panel (SHP). Even for researchers planning research on only one of the contributing countries, CNEF offers the distinct advantage of

providing a set of generated and fully imputed variables (in particular income components and equivalence weights) that are not directly available with the original surveys. These generated variables can be merged with the original survey data (and are part of the data distribution, e.g., of the SOEP). While being an excellent data source for comparative analyses of retirement and income security in old age, comparability issues have somewhat limited the use of the data for other fields such as health (the 2006 CNEF release added a set of variables that measure health behaviors and health outcomes). One disadvantage of CNEF as a multinational dataset is that the data cannot be accessed jointly and access is not harmonized across surveys.

The *Survey of Health, Ageing and Retirement in Europe (SHARE)* is a multidisciplinary and cross-national panel database of microdata on health, socio-economic status, and social and family networks of individuals aged 50 or over. Individual interviews (overall N=33,000, about 3,000 respondents in Germany) are conducted with primary respondents and cohabitating partners. Fifteen European countries including Israel have so far contributed data to the first and/or second waves of SHARE conducted in 2004 and 2006 respectively. The survey's third wave of data collection will collect retrospective life histories in sixteen countries in 2008-09. SHARE is partly harmonized ex ante with the US Health and Retirement Study (HRS) and the English Longitudinal Study on Ageing (ELSA) (see below). Data collected include health self-reports, some biomarkers, psychological variables, economic variables (current work activity, sources, and composition of wealth and current income), and social support variables (e.g., assistance and transfers within, volunteer activities). One noteworthy feature of the international SHARE database is the inclusion of anchoring vignettes for internationally comparative studies on a wide range of subjective survey data, such as health self-reports, life satisfaction, job satisfaction, satisfaction with health care, or political efficacy. Future waves of SHARE will include the collection of biomarkers and linkage with administrative records.

The *Generations and Gender Survey (GGS)* is a cross-national, multidisciplinary study of the dynamics of family relationships, covering non-institutionalized individuals aged 18-79. The main topics of the survey are related to the respondents' current and past family situation and family-related events, such as partnership formation, childbearing, and leaving the parental home. The GGS is designed as a panel survey with at least three waves at an interval of three years. GGS wave 1 data are currently available for Bulgaria, France, Georgia, Germany, Hungary, and Russia.

The *English Longitudinal Study of Ageing (ELSA)* is an interdisciplinary biannual panel survey on health, economic position, and quality of life of older adults in England. ELSA covers the range of topics needed to study the economic, social, psychological, and health elements of individual and societal ageing. The initial ELSA sample (N=12,000), interviewed in 2002,

was drawn from respondents (aged 50+) to the Health Survey for England (HSE), contributing baseline data on respondents' health (details of morbidity, lifestyle, diets, and blood samples). Covered topics are similar to SHARE: health and disability; cognitive functioning; income and wealth; employment and retirement and post-retirement activities; social networks, support, and participation. Biomarkers (blood pressure, blood samples, including genetic information, measured height and weight, lung function, grip strength, balance) are collected every four years during a separate visit by trained nurses. A life history interview has been conducted between regular waves in 2007.

The *Italian Survey on Household Income and Wealth (SHIW)* is unusual in the sense that it is conducted by a government agency (the Italian central bank – Banca D'Italia), has a strong research component, and grants free public online data access. All documentation is available in English. SHIW was started in the 1960s collecting data on income and savings. The survey has developed into a general household survey which includes detailed information on employment, wealth, financial decision making, and financial behavior.

The *Irish Longitudinal Study on Ageing (TILDA)* will interview a sample of some 10,000 individuals aged 55 and older living in the Republic of Ireland and collect detailed data on social networks and support, economic circumstances and health, including biomarkers. Three main waves of data collection are planned in 2008, 2013, and 2018, with annual telephone follow-ups and smaller-scale add-on studies. TILDA is specifically designed to deliver data comparable to the US HRS and ELSA. Data access is not yet possible but planned to be free to scientific users.

The *US Health and Retirement Study (HRS)* is the role model of many aging surveys currently conducted throughout the world, in part due to its scientific productivity. Since 1992, the HRS has generated more than 500 papers in peer-reviewed journals. Today, some 60 papers using HRS data appear in a refereed journal every year. The HRS is a biannual panel survey of Americans aged 50 and over, started in 1992. The current sample consists of about 22,000 respondents, providing detailed information on their physical and mental health, insurance coverage, financial status, family support systems, labor market status, and retirement planning. In addition to standard HRS questions, HRS has recently introduced an "enhanced" interview, collecting physical performance measures (grip strength, puff test, timed walk, balance), anthropometry (height, weight, waist), blood pressure, dried blood spots (HBA<sub>1c</sub>, cholesterol, high-density lipoprotein, C-reactive protein), even salivary DNA (for extraction and storage only).

The *Mexican Health and Aging Study (MHAS)* is a panel study of respondents aged 50 and over in Mexico. At baseline (in 2001), about 15,000 individuals have been interviewed, providing information on health (self-

reports and biomarkers), financial and time transfers between generations, sources, amounts of income, types and value of assets, and housing. The survey design closely follows HRS. MHAS provides comparable data. Respondents were reinterviewed once in 2003. New interviews are planned with MHAS wave 1 & 2 survivors and a representative sample of cohorts born 1951-1959.

The *Korean Longitudinal Study of Ageing (KLoSA)* is a biannual panel survey of approximately 10,000 South Korean residents (excluding Jeju Island) aged 45 or older, started in 2006. The 2006 main survey includes only respondents living in private households. Collected data include employment status, income, asset, family relations, health, and subjective judgment. KLoSA was designed to allow comparative studies with HRS, ELSA, or SHARE.

Based on the SHARE baseline questionnaire, the first wave of data collection for the *Japanese Study on Aging and Retirement (JSTAR)* took place in the first half of 2007. Interviews have been conducted with some 4,300 individuals from five municipalities (Takikawa, Sendai, Adachi, Kanazawa, and Shirakawa). A second wave of data collection (extending the sample to include a sixth municipality) is currently being prepared and will begin in January 2009. Although JSTAR features many design elements of SHARE and HRS, there are also a few major differences. First, the sample was drawn in only five municipalities. One of the reasons for this design is that administrative health records are stored at the municipality level. Data linkage between these records and the survey data is planned. Second, the initial sample is restricted to individuals aged 50-75. Third, JSTAR interviews only one respondent per household. However, some limited information on cohabitating spouses or partners is gathered during the interview.

*Aging studies for China, India, Thailand:* Planning for initial waves of panel surveys comparable to HRS and SHARE is underway in further Asian countries. These are the Chinese Aging and Retirement Longitudinal Study of Aging (CHARLS), the Health and Retirement Study for Thailand (HART), and the Longitudinal Survey of Aging in India (LASI).

### **3. Recommendations for future research**

This section describes selected future needs in the area of empirical research on aging. The selection reflects ongoing innovative efforts in connection with some of the surveys described in the previous section (e.g., biomarkers), but also identifies issues that have hitherto received little attention, such as the systematic inclusion of individuals living in institutions (i.e., nursing homes).

### *3.1 Data linkage*

Survey data, as described above, cover a wide range of topics. Information provided by respondents, however, is often incomplete and inaccurate. In contrast, administrative data are (ideally) complete and accurate but contain only very limited information, typically only information that is immediately related to the purpose of the data producer. The advantages of both types of data can be combined by linking administrative records to survey data. Benefits of linkage include (1) validation of respondents' self-reports, in particular if these reports are potentially subject to recall bias; (2) improved measurement of explanatory and dependent variables, reducing bias and increasing precision of model estimates; (3) reduction of respondent burden. For instance, aging surveys can benefit from adding social security records to explain retirement behavior or measurement of economic resources during retirement. Doctors' or health insurance records can be used to improve the measurement of health.

Researchers trying to link administrative data with survey data face several challenges. First, they need to get a unique ID from the respondent (e.g., social security number), which is needed to actually link the data. Asking for this ID (in some countries together with a written consent to link data) can have adverse effects on response or retention rates. Second, the availability of data that could be matched to general population surveys may be limited or require a great deal of cooperation from many agencies (such as getting medical records for German residents from public or private health insurers). Third, due to privacy legislation, data dissemination rules are often not able to conform to the standards set by the survey data to which the administrative data are linked. Restricted access to linked data will make cross-national analyses using linked data a very difficult enterprise. Here, new cross-national solutions of data access that fulfill all legal confidentiality requirements are needed.

### *3.2 Biomarkers*

One important recent development in social survey research is the integration of biomarkers. Biomarkers are often associated with genetic information such as DNA samples. The purpose of genetic research in social sciences is not to find *the* gene for some socially relevant trait. Rather, the collaboration of geneticists with social scientists is fruitful because social scientists are experts in measuring social phenomena that may or may not be associated with genes. The two important contributions of social scientists to genetic research are, first, to help to establish the importance of non-genetic factors or interactions of genes and environment, and second, their familiarity with

using large scale social, nationally representative samples to help verify associations found in small-scale medical studies.

The vast majority of biomarkers currently collected and analyzed are non-genetic: anthropometric measurements (height, weight, waist circumference, lung capacity, grip strength, balance), and blood and saliva samples. The scientific value of collecting such biomarkers in large surveys is promising: first, biomarkers improve the measurement of health. Self-reports of health are subject to considerable under-, over-, or misreporting, depending on the circumstances and dimensions at hand. Objective information can be used to validate respondents' reports and to study the amount and determinants of under-, over-, or misreporting in population surveys. However, self-reports of health have their own distinct scientific value. Thus, biomarkers should be seen as complementary measurements rather than substitutes. Second, biomarkers allow studying physiological pathways in the complex relationship between social status and health, providing information on important links that can be used to identify causal relationships. Third, biomarkers provide direct information on pre-disease pathways, in particular by measuring physiological processes that are below the individual's threshold of perception. Combined with longitudinal data on individuals, this information helps to identify the role of the environment in turning health risks into manifest diseases.

Several constraints have been identified for the collection of biomarkers in social surveys. First, collecting biomarkers increases the cost and complexity of data collection. Additional costs, like those associated with visits by trained nurses (practiced by ELSA, for example), may seem prohibitive. Recent developments in minimally invasive methods which allow collection by trained survey interviewers have led to considerable cost decreases. For instance, it is now possible to measure HbA1c (as a measure for diabetes), cholesterol (to measure risk of cardiovascular disease) or C-reactive protein (to measure risk of cardiovascular disease and chronic stress) in dried blood spots. Thus, the most important risk factors for chronic disease and work disability in early old age can be measured through the collection of a few drops of blood taken from the finger tips. Together with measuring blood pressure and tobacco consumption, these data allow forecasting for the incidence of cardiovascular disease 10 years hence. Thanks to advances in technology, the costs of collecting blood and analyzing the samples in laboratories are now down to a few Euro per respondent.

Still, while the research potential of collected biomarkers is large, the training of lay interviewers and the logistics of storing and sending specimens should not be underestimated. Second, collecting biomarkers increases respondent burden and may affect the willingness of survey participants to cooperate in future waves. Third, biomarkers are potentially sensitive information and raise a lot of ethical issues surrounding confi-



dentiality, storage, and respondent information. Survey researchers planning to include biomarkers in their data collection efforts need to be aware of these constraints.

### *3.3 Coverage of nursing home residents*

In the US, some 15 percent of individuals aged 85 or older live in nursing homes. Thirty percent of all individuals die in nursing homes. However, the social and health determinants of nursing home admissions and the living conditions and quality of life of nursing home residents is greatly under-researched. The main problem faced by survey researchers probably is that in many countries, including Germany, no sample frames exist that include reliable information on nursing home residents (or other institutionalized populations). The typical approach taken so far is to draw a baseline sample from the non-institutionalized population and to follow respondents who move into nursing homes between waves. In principle, this approach should lead to samples of nursing home residents of reasonable size if panel surveys mature. In practice, however, there is substantial under-coverage due to the problems involved with tracking respondents, gaining access, and also due to a lack of respondents' ability to answer (and an increased need for proxy respondents). The oldest-old are the fastest growing segment of the population, and dementia – already a leading cause of nursing home admissions – is likely to be an increasing concern among the oldest-old. Despite the challenges of collecting data on them, neglecting a significant proportion of the older population in social surveys is hard to justify. Recent experience, for example from the Danish Longitudinal Centenarian Study, shows that many concerns voiced about conducting interviews in nursing homes (e.g., unethical, too costly) are unfounded. The primary impediment in most countries is the lack of a suitable sample frame such as a nationwide person register.

## **4. Conclusions**

The collection of multidisciplinary, longitudinal data on aging is one of today's most active and innovative fields in survey data collection. This report has documented the host of data available to researchers from various disciplines working in the field of individual and population aging. We can observe several exciting developments. First, researchers across different countries (including emerging economies) are trying to collect data that are comparable internationally. Comparability is sought mainly with US surveys, because the US not only plays a leading role in survey methodology, but also

because it is an important reference country, due to its size and due to the particularities of its welfare state. Second, data access for secondary analysis has become easy and quick, thanks to developments in information technology. With few exceptions, data are released after some cleaning often less than one year after they were collected. For many surveys, released data are downloadable from websites literally within minutes. Thus, the concepts of primary and secondary analysis become meaningless. Rather, it makes more sense to speak of overlapping groups of data producers and data users. Third, researchers are currently trying hard to bridge boundaries between disciplines, especially between social sciences and medicine. Some biomarkers are already routinely included in a number of ongoing surveys, and the scope of measures that can be collected during normal face-to-face interviews is increasing due to technical progress.

## 5. Appendix: Summary information on current aging surveys

Survey	Country	Age range	Sample size*	Survey Years
SOEP	Germany	17+	22,000	1985-2007
DEAS	Germany	40-85	6,000	1996, 2002, 2008
SAVE	Germany	18+	3,000	2001-2007
CNEF	USA, Germany, UK, Australia, Switzerland, Canada	17+	181,000	1980-2007
SHARE	Denmark, Sweden, Austria, France, Germany, Switzerland, Belgium, Netherlands, Spain, Italy, Greece, Poland, Czechia, Israel, Ireland	50+	35,000	2004-2008 (biannual)
GGG	Bulgaria, France, Georgia, Germany, Hungary, Russia	18-79	60,000	2005, 2008
ELSA	England	50+	12,000	2002-2008 (biannual)
SHIW	Italy	18+	20,000	1977-2006
TILDA	Ireland	55+	10,000	2008
HRS	USA	50+	22,000	1992-2008 (biannual)
MHAS	Mexico	50+	15,000	2001,2003
KLoSA	Korea	45+	10,000	2006,2008
JSTAR	Japan	50-75	4,300	2007,2009

## Access

SOEP data are available for all academic users from DIW upon signature of a user contract.

Access to the first two waves of DEAS data is provided via the central archive for social science data (ZA) at the University of Cologne (Study Numbers: 3264, 4304), access class C: data access is granted to academic users upon approval of the primary researchers.

SAVE data are available through the central archive for social science data (ZA) in Cologne (Study-Numbers: 4051, 4436, 4437, 4521, 4740), access class C: data access is granted to academic users upon approval of the primary researchers.

The PSID-CNEF file is public use and can be simply downloaded from the CNEF website. Access to BHPS-CNEF, SOEP-CNEF, or HILDA-CNEF requires approval by the BHPS, SOEP, and HILDA primary researchers, respectively. SLID data (a non-research driven survey conducted by Statistics Canada) can only be accessed via remote computing.

SHARE data are available online to academic users via the SHARE website (upon signature of a data confidentiality statement) or through the central archive for social science data (Zentralarchiv) in Cologne (Study-Number: 4560), access class C: data access is granted to academic users upon approval of the primary researchers.

Data access is granted only after a research proposal submitted to the data administrators has received a positive review "for relevance to research". Who the reviewers are and what criteria for relevance they use is as yet unclear.

Online access to Scientific Use Files for waves 0 (i.e. HSE data) through 3 (i.e. the 2006 data collection) is available via the UK Economic and Social Data Service (Study-Number 5050).

Microdata are freely available from the Banca D'Italia website.

Data are not yet available.

Original HRS data are available to researchers after a simple online registration process.

Additionally, a user friendly combined and harmonized HRS file is made available by RAND.

Data collected so far are freely available from the MHAS website after a simple registration process.

The 2006 data and English documentation are freely available from the KLoSA website after a simple registration process.

Public release of wave 1 data planned for 2009.

\* Note: Sample sizes may vary from year to year.



## 5.7 Income Provisions and Retirement in Old Age

*Tatjana Mika, Uwe Rehfeld and Michael Stegmann*

Contact:

Tatjana Mika  
Federal German Pension Fund  
Deutsche Rentenversicherung Bund  
Research Data Center of the German Pension Insurance  
Ruhrstraße 2  
10709 Berlin  
Germany  
e-mail: tatjana.mika[at]drv-bund.de  
uwe.rehfeld[at]drv-bund.de  
michael.stegmann[at]drv-bund.de

## **Abstract**

Research on the income situation of current and future retirees often requires record-based data. Because records include accurate information on the life course, they can also, when linked to survey data, make interviews shorter and less demanding for the interviewed persons. Process-produced data from the German Pension Insurance are already available for research topics in this area. The data include details about employment histories and other life-course events insofar as they are considered in the calculation of pensions. Nevertheless, additional sources are needed if research projects are to address the income situation in more detail, in particular the question of poverty or affluence in old age. The pension reforms of the past decade have strengthened the importance of the second and third pillar, thereby increasing the importance of occupational pensions and private savings for future old age income. There are already some detailed and inclusive data for research on old age income and retirement that have been collected for government reports, but not all this data is available for scientific research yet. Furthermore, the exchange of data between social security and/or tax institutions should be combined more often with the collection of statistical data in order to improve the possibility of record-to-record linkage.

Keywords: retirement, old age provisions, public pension fund, process-produced data, data linkage

## **1. Research questions**

A principal theme in current research on old age provision is that of income after retirement for current and future retirees. For an analysis of the current retired generation comprehensive data on all income sources are needed. For in-depth research on the reasons for specific income situations among retired persons, the data should include life-course information on employment and income over the life cycle until retirement. Process-produced data from the German Pension Insurance are especially useful for these research topics because the public pension scheme is the most inclusive old age provision and it also contributes the largest share of income received in old age. These data include details about employment history and other life-course events insofar as they are considered for the calculation of pensions. The life-course information collected is very broad because of the far-reaching evaluation of social situations and activities dictated by past and current pension law. Nevertheless, additional sources are needed if research projects are to address the income situation in more detail, in particular if they are to answer questions about poverty or affluence during old age. Occupational pensions and private savings are important additional components in old age provisions. The pension reforms of the past decade have strengthened the im-

portance of the second and third pillars of occupational and private pensions, thereby increasing their importance for future old age income. The generally lower income of women in the older population, especially in West Germany, also requires information on household income in order to assess the real economic situation.

The central focus of social research is the age of retirement, which is largely determined by social security law. The pension law determines overall retirement behavior among those who are socially insured. Process-produced data are the best choice for this kind of research question, including information on legal background that is unknown even to the pensioners themselves. However, process-produced data do not include subjective information on the motives behind early retirement. The data are not accurate for persons who retired earlier from lifetime civil service employment or from being self-employed, but who only received their public pension later. As a result, pension insurance data alone lead to an overestimation of retirement age.

To assess future old age income requires a different approach and more adequate data. To begin with, the forecast of future income requires a thorough collection of all information about old age provisions undertaken up to the date of retirement; including acquired social security rights, private insurance, and other savings and occupational pensions. These data then form the basis for estimates of future old age income.

## **2. Databases and data access**

### *2.1 Databases for current old age income*

Data on the income situation of the current aged population are available in many surveys like the German Socio-Economic Panel (SOEP, *Sozio-oekonomisches Panel*), the Microcensus, and the EVS (*Einkommens- und Verbrauchsstichprobe*). The varying income structure of retired people nevertheless requires a special survey design. Difficulties may arise because persons over a certain age are often difficult to interview and can be hard to reach if they live in an institution and no longer in a private household. Process-produced data are therefore helpful to estimate the number of elder people who are not or no longer available for surveys. They also offer information about the legal conditions of a granted pension.



### 2.1.1 Pension Records (RTBN, *Rentenbestand*) and Completed Insured Life Courses (VVL, *Vollendete Versichertenleben*) samples

The sample of the pension records includes all pensions paid from the German Pension Insurance at the last day of each year. These data are a useful basis for the validation of other sources on the retired segment of the population. Participation in the public pension insurance scheme is mandatory for all persons in Germany employed in the private or public sector. Additionally, contributions are paid out of unemployment insurance in the case of the unemployed, out of health insurance in the event of long-term illness, and from the state for people in military or civilian national service. The majority of the population thus comes into contact with the pension insurance system at some point or another in life, and the pension insurance system has data on about 90 percent of the entire population. The statutory old age and disability pension – due to its income replacement tasks and broad social basis – provides the main income source after retirement. Survivor pensions are the main source of income for widowed women.

The special survey, Completed Insured Life Courses, (VVL 2004) is a useful source for empirical analysis about retirement age and income in relation to the life course. A 20 percent sample of newly granted pensions in a particular year is the basis for this longitudinal data. The calculation of the pension is one important source of information, in combination with the longitudinal dimension of information about the past – from age fourteen until retirement. The sampling of data from one year of newly granted pensions enables researchers to compare different life courses, ending in the social status of becoming a pensioner in the same year in East and West Germany and abroad. At the time of retirement, the pension fund has gathered all information on a life course as far as the activities, contributions, and legal entitlements are relevant to the pension benefit. The moment of retirement is the point where people hand over all necessary proofs to the pension insurance office in order to receive a pension on the basis of all the relevant facts. From the statistical point of view, it is therefore the point in time at which information about the life course is most accurate. The sample drawn from all newly granted pensions is so large that all social strata and many different types of life courses are represented in sufficiently large numbers to enable empirical research on many different questions. Only old age and disability pensions are selected for this sample, excluding survivors pensions.

For the dataset “Completed Insured Life Courses,” all pension funds send the information on the completed biography – from secondary school to the moment of retirement – to the Research Data Center of the German Pension Insurance (RV, *Deutsche Rentenversicherung*) (Stegmann 2007). The Scientific Use File for social research combines the longitudinal life-

course information with the result of the pension calculation as a cross-sectional part of the dataset. The cross-sectional part includes additional demographic information. This means that demographic variables mirror the social situation at retirement. However, in some cases, the socio-demographic position might have changed over the lifetime, for example, if immigrants with foreign nationality are naturalized at retirement or if marital status changes over the lifetime. The longitudinal information is presented on a monthly basis. For each month the data shows whether the person was gainfully employed or was in another social position such as unemployment, care-giving, or sickness. Childcare is assumed to be the main occupation if the birth of a child is registered and no gainful employment has taken place afterwards.<sup>1</sup> Employment has priority status in the data and all other social situations are second in rank. A lack of information means that a person is in none of these social status situations at this time in Germany. Such a gap in information can stand for self-employment without social insurance obligation, unemployment without being entitled to benefits from the Federal Employment Agency, or working abroad. However, in most female biographies a gap in information stands for a period of housekeeping.<sup>2</sup> The main drawback of these data is the lack of information on other sources of income before and after retirement.

The data are accessible at the Research Data Center of the German Pension Insurance. A smaller sample is drawn for the Scientific Use File, which can be ordered for use in research institutions. Larger samples up to full samples in the case of data on recorded pensions can be use on-site at the Research Data Center.

### 2.1.2 Survey on Old-age Pension Schemes in Germany (ASID)

It is the aim of the research project called Old-age Pension Schemes in Germany (ASID, *Alterssicherung in Deutschland*) to provide up-to-date and representative data organized according to various socio-demographic groups that describes the income situation of the older population in West and East Germany. The first survey was realized in 1986, the last and current study was carried out in 2007. The law requires that the ASID survey be conducted once every legislative period because it is the basis for an official government report on the income of the older population (*Alterssicherungsbericht*). The study has so far been carried out six times. The population studied

---

1 The birth of a child is registered in the pension record of one of the parents. In most cases this is the mother, because there is an income cap that hinders higher earners from profiting from the child benefit. The child benefit in the German Pension Insurance credits one point for children born before 1992 and 3 points for children born after 1992.

2 This fact can be proven with the data from the AVID (*Altersvorsorge in Deutschland*) 1996 project, where process-produced data were combined with survey data.

includes all those older than fifty-five, including those living in residential homes.

The centerpiece of the survey is the collection of data on over twenty-five types of income. In the case of couples the data include the income of both spouses and in the case of widows income is disaggregated into self-acquired and derived benefits. Like all old age pension provision systems, the ASID collects information on income from various sources, but also includes income from private sources like private insurance. Thus, the ASID shows not only the level of overall gross and net income but also allows for varied analyses of income patterns. This is supplemented by information about the course of working life and the current life situation. In the case of couples, relevant data is acquired for both spouses, in case of widows the late husband is also taken into consideration. The gross incomes are finally converted into net incomes by way of an income tax and social insurance contribution model.

For the ASID, a representative sample is taken for this particular purpose from the local registers, where all people living in Germany must register. This is the best and most expensive method of sampling survey data. It can only be used if there is a public interest in the conducted survey. Irrespective of marital status, both men and single (i.e., widowed, divorced, and unmarried) women were chosen at random as target persons. The data of married women were collected together with the data gained from their spouses. The survey thus includes married couples and single persons. The data of the ASID 1992, 1995, and 1999 can be ordered via the GESIS Data Archive for the Social Sciences (*Datenarchiv für Sozialwissenschaften des Leibniz-Instituts für Sozialwissenschaften*) in Cologne using the keyword “Alterssicherung in Deutschland.” For later surveys, the data have not been published and a date of publication has not yet been announced.

## 2.2 *Old age provisions of future pensioners*

The estimation of future old age income on the basis of current acquired pension rights and savings is a particularly daunting task. It requires the assessment of contributions to public, occupational, and private schemes up to the present and the prospect of future old age income streaming from these sources.

### 2.2.1 Sample of the insured population pension fund records

The sample of all actively insured persons in a specific year, called the insurance account sample (VSKT, *Versichertenkontenstichprobe*), is the best source for acquired rights in the public pension scheme. The longitudinal

information is presented in the same format as in the “Completed Insured Life Course” data, but the former sample includes 500,000 people from the insured population aged seventeen to sixty-seven. Insured persons with non-German citizenship are over-sampled because they have a high percentage of incomplete records due to migration. In this case, transnational mobility severely limits the collection of complete life-course information.

The social situations of persons sampled are recoded from the original mainframe data into similar social situations consistent with the “Completed Insured Life Courses” survey described above. Future old age income is estimated in the data on the basis of contributions paid up to the time of sampling. The estimated pension included in the data is calculated as if the insured person retired with a disability pension for health reasons at the date of sampling. All information used for the calculation of a pension from the public pension scheme is also included in this estimate. This includes raising children, periods of education and training, as well as phases of unemployment and care-giving. The sample size and accurate information on employment and socially insured gainful employment make the VSKT attractive for social research, but the lack of information on other sources of income and the household income are a drawback for research on future old age income. A selectivity problem arises insofar as people who were employed as lifetime civil servants after less than five years of socially insured employment are not registered. Self-employed professionals have, on the other hand, a very incomplete socially insured life course, but are often covered by other compulsory old age provision schemes. The selectivity problem could only be undone if all compulsory old age provision schemes would send their data to a common statistical collection point. For occupational pension schemes data are still lacking, but the Federal Statistical Office is conducting a feasibility study about the future collection of data on this retirement income source.

#### 2.2.2 Combination process generated with survey data from Retirement Pension Provision Schemes in Germany

The last decades have shown a diversification in and growing heterogeneity of (empirical) life courses in Germany as well as in many other welfare states. Facing this evolution, the German Pension Insurance and the Federal Ministry for Labor and Social Affairs (BMAS, *Bundesministerium für Arbeit und Soziales*) commissioned an elaborate study on future old age incomes with a strong biographical focus called the Retirement Pension Provision Schemes in Germany (AVID, *Altersvorsorge in Deutschland*).

The target persons of the first AVID Study (AVID 1996) stem from a representative random sample drawn from the panel of samples of statutory pension insurance accounts. A tailor-made projection of the data guarantees that the outcome of analyses of those contributors to the statutory pension

scheme aged forty to sixty and their spouses will have a representative character. The projection method chosen makes both single person and spouse-related evaluation possible. The universe of analysis of the AVID 2005 was extended to include the entire population of Germany born between 1942 and 1961 (and their spouses), irrespective of any entitlements to pensions from the German Pension Insurance. As a result of this decision, the representative sample had to be derived from a different source and was in this instance taken from an access panel. The record to survey data match of the first wave had the advantage that the quality of the pension insurance data was the same as for the "Sample of the Insured Population Records." The price for this was that the data were representative only for the population who had a pension insurance account on the date of sampling. The second survey to record paths ensured representativeness for the whole population, but created a more troublesome process in the collection of process-produced data afterwards (Frommert and Heien 2006; Bieber and Stegmann 2002).

The objective of this project is to identify the type and amount of entitlements to old age income for individuals and married couples, that is, for pension-insured persons between forty and under sixty years of age (age groups born between 1936 and 1955, Germans living in Germany and – irrespective of nationality and age – their spouses). Thus, in AVID 1996 for the first time entitlements to payments from the statutory pension insurance for those between forty and under sixty years of age among married couples are shown, the accumulation of entitlements within the statutory pension insurance are covered (including entitlements arising from other standard and supplementary pension systems), previously unavailable information in terms of gaps in the insurance biographies of the statutory pension insurance is gathered, and extrapolation data are supplied for the analysis of future developments. Moreover, various data concerning life and working biographies are supplied that are not included in the individual pension accounts.

The AVID studies are characterized by an innovative mix of methods and data sources. They are composed of several steps, each of these and their combination as a whole are rather unique in the context of old age security research.

The survey comprises all the important schemes for retirement in Germany. Apart from the German Pension Insurance, the most significant system, this includes private and public supplementary systems, the civil servants' pension scheme, farmers' old age pensions, and schemes for independent professions. Private provisions such as life insurance and private pension insurance, ownership of property, as well as maintenance payments made by children and partners are also taken into account.

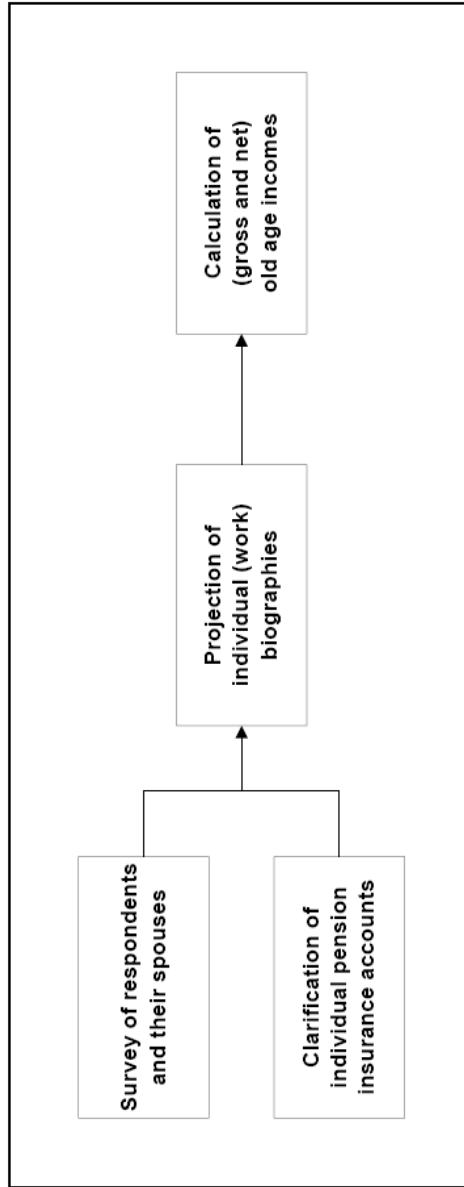
After the deduction of income tax and statutory contributions for health and old age insurance, the net old age income shown is the sum of benefits

due to personal entitlements arising from the pension schemes described above, including private provisions such as life insurance and private pension insurance, survivor income, or other benefits, if any, arising from such schemes. Other sources of income such as earned income or transfer payments (i.e., housing subsidies, welfare benefits, and other forms of unearned income) are not taken into account. In the second step, the individual pension insurance accounts of the AVID respondents are clarified by the (federal or regional) institutions that manage the accounts; the entire process takes about eighteen months. These two datasets are then matched with the respondents' consent and checked extensively for consistency, resulting in a highly valid and reliable dataset on (past) life courses and pension provisions. The design of the survey also allows for the identification of married couples so that at least some measure of household context can be included in the analyses. In the third step, the individual (work) biographies are projected to the age of 65 – at the time of the surveys the legal retirement age in Germany – using a specially developed micro simulation model. Biographical events like unemployment and long-term illness are taken into account as well as individual decisions to interrupt or end employment for purposes such as housekeeping, raising children, or looking after relatives in need of care. The projection is based on a projection corridor using individual data on the years 1992 to 1996 for the AVID 1996; the projection corridor for the new study is correspondingly longer (1992 to 2002). The simulation model does not take into account socio-demographic processes, so marital status represents the status of the survey year and any mortality is excluded.

The findings are based on the projected old age income that is calculated in the final step: the gross old age income at the age of sixty-five is calculated on the basis of individual biographies, taking all the relevant pension schemes – the statutory pension insurance, the civil servants' pension scheme, the farmers' old age pension scheme, special schemes for the liberal professions, the public and private supplementary systems and private provisions for old age (life insurance, private pension insurance) – into account. Finally, the net incomes are generated by taking into account current income tax regulations and a specially designed social insurance contribution model.

The results of the AVID Studies are published as reports and charts. The data have not been published for the scientific community. However, due to its combination of process-produced data from the records of the federal pension insurance and survey data, the AVID sets an example for future data based on survey-record linkage.

Data sources of the AVID:



Source: Dina Frommert und Thorsten Heien (2006)

### **3. Future developments**

#### *3.1 Record-to-record linkage*

Record-to-record linkage would improve data where other social security or tax institutions gather information unknown to the pension fund. In the case of the Federal Employment Agency, there are relevant data on training and other benefits from unemployment insurance. In the case of public health insurance, such data would include information about prescriptions and health treatments.

Record-to-record linkage is not an easy task for the Research Data Center of the German Pension Insurance (RV, *Deutsche Rentenversicherung*), because the social security number is not known for persons included in the data gathered for statistical purposes which is the case, for example, for the 500,000 persons included in the insurance account sample (VSKT). The public pension fund itself is a federation of several regional and two federal insurance programs that keep the records on the persons insured by them. For statistical purposes all pension insurers send data to the central statistical collecting point, not including the social security number. The Research Data Center is therefore unable to re-identify the persons included in the data provided for scientific research. Record-to-record linkage must therefore be supported from all pension insurers by sending the data including the social security number for a particular research project and a particular year and data source. These projects must therefore be presented to the self-governing boards. Strict data privacy rules also apply because the Research Data Center is not allowed to have the social security number in its reach. A regular procedure is not yet in place for record-to-record-linkage, but if convincing new results from research projects using improved data were presented, a recurrent regular procedure could become possible. The next step in record-to-record linkage would be matching data that have other identifiers other than social security numbers. These could be tax data or data from the old age provision schemes.

#### *3.2 Matching process-generated data with survey data*

The linkage of pension insurance records with survey data would improve the life-course information for many surveys, which suffer from recollection errors by interviewed persons or left censoring. Research on retirement or disability would, on the other hand, gain from subjective information gathered via the survey. Self-assessed health and retirement planning would give crucial insights about early retirement, a research topic of the utmost importance. Objective indicators of health status could also be included (biomarkers). However, for survey-to-record linkage re-identification is likewise



not an easy task. The self-governing boards of the pension insurances have to give their consent to the project, because the cooperation of all pension insurances is here again essential. The survey must include the informed consent of the interviewed person, who must also provide his or her social security number. The collected number must then be searched for in the central register of all socially insured persons. The next step is the collection of the data from pension insurance records.

#### **4. Future developments: European and international challenges**

The German Pension Insurance exchanges data with most public pension funds worldwide because social security treaties require that migrants should not be discriminated against and should have an easy one-step procedure to apply for their pension in just one country. The most important partners in data exchange are the Member States of the European Union, who act under the common framework of the same regulations. The exchanged information is only minimally included in the statistics. A common effort to gather more information on transnational working biographies would be a step toward improving the data.

A common international pool of data on public or publicly supported old age provisions would promote comparative research. This would require improved multilingual metadata and very accurate documentation, because the differences between national security systems could easily lead to misconceptions.

#### **5. Conclusions and recommendations**

Research on the income situation of current and future retirees often requires record-based data. Because of their accuracy in providing life-course data they also can, if they are linked to survey data, make interviews shorter and less demanding for the interviewed persons. However, the projects leading to this improved data are time-consuming, because privacy laws demand extensive data protection requirements. Furthermore, social security laws require that the collected data are for research projects on social security related topics like health status and early retirement. Record-to-record linkage does not require informed consent to protect privacy, but extensive technical provisions separate the data from the identifier at an early stage. Both

methods of matching greatly improve data quality at a reasonable price, but require an established infrastructure that is able to handle the procedures required by privacy laws and regulations.

## References:

- Bieber, U. and Klebula, D. (2001): Kräftiger Aufholprozess: ASID 1999. Bundesarbeitsblatt 9/2001.
- Bieber, U. and Stegmann, M. (2002): Sozialer Ausgleich in der gesetzlichen Rentenversicherung. Empirische Ergebnisse auf Basis der Untersuchung AVID 1996. Deutsche Rentenversicherung 11/2002.
- Federal Ministry of Labor and Social Affairs (2001): Old Age Pension Systems in Germany 1999. Summary of results.
- Federal Republic of Germany (2005): National Strategy Report on Old-Age Pension Provision 2005 (NSR 2005).
- Frommert, D. and Heien, Th. (2006): Retirement Pension Provision Schemes in Germany 1996 and 2005. Schmollers Jahrbuch 126 (2), 329-336.
- Kortmann, K. and Schatz, Ch. (1999): Altersvorsorge in Deutschland 1996 – (AVID'96). Zusammenfassung wichtiger Ergebnisse der Untersuchung „Strukturen und Trends der Altersvorsorge von 40- bis 60-jährigen Rentenversicherten und ihrer Ehepartner“. DRV 10-11/1999, 573-597.
- Roth, M./Stegmann, M. and Bieber, U. (2002): Die Aktualisierung der Studie Altersvorsorge in Deutschland – Inhaltliche und methodische Neuerungen der AVID 2002. DRV 11/2002, 612-641.
- Stegmann, M. (2007): Biografiedaten der Rentenversicherung und die Aufbereitung der Sondererhebung „Vollendete Versichertenleben 2004“ als Scientific Use File. DRV-Schriften Band 55/2006, 82-95.
- TNS Infratest Sozialforschung (2005): Alterssicherung in Deutschland 2003 (ASID '03) – Zusammenfassung wichtiger Untersuchungsergebnisse, Forschungsbericht Bd. 346/Z des Bundesministeriums für Arbeit und Soziales. Berlin.
- TNS Infratest Sozialforschung (2007): Altersvorsorge in Deutschland (AVID) 2005: Alterseinkommen und Biografie. Forschungsprojekt im Auftrag der Deutschen Rentenversicherung Bund und des Bundesministeriums für Arbeit und Soziales. Berlin.
- Verband Deutscher Rentenversicherungsträger (VDR) (2000): Special Survey „Old Age Pension Provision in Germany (AVID'96)“ Old Age Provision of Statutory Pension Insurance Contributors – Birth Years 1936 to 1955. Übersetzung des Projektdesigns. VDR. Frankfurt/M.



6. POLITICAL AND CULTURAL PARTICIPATION AND  
THE ROLE OF THE MEDIA

6.1 Political Participation – National Election Study

*Rüdiger Schmitt-Beck*

Contact:

Rüdiger Schmitt-Beck  
University of Mannheim  
Chair of Political Science I (Political Sociology)  
68131 Mannheim  
Germany  
e-mail: [schmitt-beck\[at\]uni-mannheim.de](mailto:schmitt-beck[at]uni-mannheim.de)

## Abstract

This advisory report gives an overview of recent developments in political participation and electoral research, and discusses the current state of affairs with regard to data provision and access. It concludes with several recommendations: (a) to retain a small number of key political variables in the future German General Social Survey (ALLBUS, *Allgemeine Bevölkerungsumfrage der Sozialwissenschaften*) and the German Socio-Economic Panel (SOEP, *Sozio-oekonomisches Panel*) questionnaires to create substantial amounts of synergy at little marginal cost; (b) to establish a National Election Study in Germany by providing the current German Longitudinal Election Study (GLES) project<sup>1</sup> with a constant logistic and methodological support infrastructure through the Leibniz Institute for the Social Sciences (GESIS, *Leibniz-Institut für Sozialwissenschaften*). It is recommended that in the long run a regular follow-up study to this project be provided with stable public funding and a firm institutional embedding, preferably by including it in the remit of GESIS; (c) to align the data services of statistical offices more closely to the data requirements of participation and electoral research; (d) to establish a formal obligation for public agencies to submit survey data collected under their auspices in due time to the public domain for purposes of secondary analysis.

Keywords: political participation, political behavior, elections, electoral behavior, voting

## 1. Introduction

The notion of political participation, in the sense of voluntary activities undertaken by free and equal citizens to influence the course of government, is at the heart of the idea of representative democracy (Dahl 1972). To be sure, in liberal democracies no one is obliged to take part in politics. But if large majorities of the citizenry abstained from any political involvement, there could simply be no democratic politics. Hence, a substantial amount of political activity on the citizens' part is essential for the functioning of democracy. Therefore, describing and explaining how people participate in politics is a vitally important task for political scientists. Consequently, patterns and dimensions of political participation, which encompass a whole range of activities, from contacting local officials to engaging in acts of political violence, have been scrutinized extensively since the 1960s (van Deth 2003). Among the many avenues by which people can make their needs and interests count in political decision-making, casting votes at general elections has always been the most important one. To the present day, it is by

---

1 The GLES is funded by the German Research Foundation (DFG, Deutsche Forschungsgemeinschaft) to study the 2009, 2013, and 2017 national elections.

far the most widely used and most egalitarian form of political action. Moreover, it stands out as the one form of political participation that, by its very nature, is inextricably tied to the core principle of representative democracy. It decides who is granted access to public office and thus to the levers of power. It is a sharp weapon in the hands of citizens that enables them to hold officials accountable to the will of the people. It seems just, then, that electoral behavior is one of the most intensely explored political phenomena. Most of this research has concentrated on explaining citizens' vote choices, while studies about turnout and its preconditions are less numerous (Falter and Schoen 2005; Lewis-Beck et al. 2008).

Seven years after the 2001 KVI report, the present advisory report will attempt to take stock of the current state of data provision and access with regard to the various aspects of political participation and, in particular, electoral behavior in Germany. It first gives an overview of recent theoretical and methodological developments within the field of participation research, especially electoral research, that appear particularly important from the perspective of data provision and access. It will then discuss the current state of affairs with regard to these two foci in Germany, including developments that have taken place since the 2001 KVI report. The chapter concludes with a brief summary of recommendations, directed at either policy-makers or scientific infrastructure organizations.

## **2. Recent developments in participation and electoral research**

Since Milbrath's (1966) seminal study on political participation in the United States in the 1960s, participation studies have flourished, and quickly developed into a respected subfield of political research. In the 1970s, the first large-scale internationally comparative projects were undertaken, and comparative survey research has, to the present day, remained the hallmark of this strand of studies. Between them, these studies have greatly enhanced our understanding of political participation – the incidence of its various forms, its dimensionality, and its backgrounds, i.e., the factors that facilitate or impede citizens' active involvement in politics (van Deth 2003; Kaase 2007). Since the 1990s, the field of participation studies has expanded and become part of a broader paradigm of research into modern democratic citizenship. This recent conception conceives political participation as one of but a whole range of facets of citizens' orientations towards their political system, including social participation (such as associational membership and activity; see the chapter on civil society by Alscher and Priller in this publication), socio-political norms and values (such as civic obligations,



tolerance, norms of reciprocity, or inclusion/exclusion); and support for democracy and its institutions (Pattie et al. 2004; van Deth et al. 2007).

The special field of electoral research has also substantially expanded its scope in several ways. Traditionally, it has been guided by a small set of related questions: who votes, and for what reasons? Which candidates and/or parties are chosen, and, again, for what reasons? Typically, these questions were focused at particular national elections. Representative surveys of voters (often cross-sections, sometimes short-term panels) were the method of choice to answer these questions. In recent projects, this rather narrow frame of surveying and collecting data has given way to a broader perspective that seeks to understand elections as part of broader processes of political representation. This includes manifold and dynamic interactions between citizens and office-holders, as well as candidates for electoral office, with political parties and the mass media functioning as mediating agencies. The broadening focus of electoral research has also been accompanied by a pronounced interest in the dynamics of the communicative processes taking place between citizens on the one hand, and parties and their candidates on the other. An implication of this is that study designs have moved from a cross-sectional to a longitudinal approach (Romer et al. 2006). Further, this shift in electoral research has forced researchers to go far beyond mere voter surveys in data collection. Consequently, researchers have also added candidate surveys, party campaign studies, media content analyses, and contextual data. Moreover, electoral studies are even coming to see inter-election periods as similarly important for election outcomes (Güllner et al. 2005). This development has heightened the increasing data requirements of studies not only during the few weeks of the “hot” campaigns immediately preceding elections, but also at more or less close intervals during entire electoral cycles. As voters’ political behavior tends to become more and more individualized and volatile, it seems clear that ideal designs to study contemporary elections need to include specific components for capturing short-term campaign dynamics immediately preceding elections, on the one hand, and for tracking the long-term changes that take place over whole electoral cycles, on the other.

Closely connected to this is a recent tendency of electoral studies to become less sociological and more political. Traditionally, election studies tended to see individual voters and their attributes as the sole key to understanding the outcomes of elections – as if these were occurring in a political vacuum. Recent studies, in contrast, try to explore how elections can be better understood by taking into account the institutional and situational political contexts within which they take place. These typically include the behavior of parties, candidates, the media, and other actors. Naturally, such a perspective would require attention to be directed beyond individual elections by comparing various elections in both cross-national and longi-

tudinal perspectives. Hence, elections themselves become units of observation in complex longitudinal and multilevel research designs (Franklin and Wlezien 2002; Thomassen 2005). Obviously, such studies are far more demanding than traditional ones in terms of data requirements.

Although older than participation studies, electoral studies lag behind this field with regard to internationally comparative projects – for obvious reasons. National elections are in many respects idiosyncratic affairs (beginning with their dates), and studying them in an internationally comparative perspective poses serious challenges in terms of study designs and instrumentation. Recent years have seen the development of measures which may be used to successfully deal with these problems. One is the Comparative Study of Electoral Systems (CSES), a collaborative program of research among election study teams from several countries around the world, including Germany. These study teams include a common module of survey questions in their own post-election studies which are further enriched with system-specific macro variables to allow for multilevel analyses and the study of interactions between system characteristics and individual behavior at elections (Klingemann 2009).<sup>2</sup> Another is the European Voter Project (Thomassen 2005), which, jointly with the GESIS Data Archive, successfully undertook the formidable task of harmonizing data from national election studies from six countries over more than four decades (Mochmann et al. 1998). It has additionally started a successor project, entitled COST Action “The True European Voter.”<sup>3</sup>

Of particular relevance for this advisory report is yet another recent trend: a palpable strengthening of interest within the political science community to move beyond single election projects and engage in creating permanent, integrated data infrastructures for electoral research. More and more countries are institutionalizing National Election Studies as part of their social science data infrastructure. In Germany, a concerted attempt to establish such a study started in 2007, which is described in more detail below. Teams of French and Austrian political scientists are also engaged in similar activities in their respective countries. It also deserves mention that a multinational team has been awarded funding under the EU’s Seventh Framework Programme to carry out a pilot study for the creation of an extensive European data infrastructure for research into citizenship, political participation, and electoral democracy at the level of the EU.<sup>4</sup> Moving beyond an exclusive emphasis on surveying voters, this project impressively illustrates the trend towards broadening the scope of election studies towards dynamic studies of political representation. Importantly, such data infrastructures are not intended to serve exclusively the data requirements of

---

2 <http://www.umich.edu/~cses/>

3 <http://true-european-voter.eu/>

4 <http://www.piredeu.eu/>

scientists specializing in electoral research, but also to address – by appropriate means of data dissemination – the information needs of the general public. This includes political actors ranging from MPs, government agencies, parties, and organized interests to journalists and members of civil society.

### **3. Data provision and access**

The 2001 KVI report did not include a special section on political participation, but it did contain an excellent, highly detailed stock-taking of provision and access to data on elections and political parties (Niedermayer 2001). With regard to elections, this report evaluated the availability of data for purposes of scientific research on the whole quite positively, although it also emphasized – to adopt Lipset and Rokkan's (1967, p. 50) famous phrase – a “few but significant exceptions” to this. One of the most significant gaps mentioned concerned the general dearth of data concerning elections at the local level. This bleak state of affairs has remained virtually unchanged. In stark contrast to European, national, and *Länder* elections, local elections have remained a “blind spot” and are therefore still extremely difficult to analyze.

Official electoral data are reliable and can therefore be used as benchmarks for data collected by means of sample surveys. Importantly, some research problems can only be addressed using this kind of data. These research problems typically include analyses aimed at understanding how political behavior is embedded in broader socio-spatial contexts, which requires advanced methods of multilevel analysis. While provision of data from official electoral statistics is generally satisfactory, from the perspective of electoral research, revisions of current practices seem desirable with regard to a number of details. One concerns the residual category of “other” parties. As a matter of information efficiency, it seems appropriate to use such condensed categories in official publications, but the results of these parties should, as a rule, always be reported separately in computerized data collections. In an age of ongoing party system fragmentation, from the perspective of electoral research, it is desirable to obtain easier access not only to data pertaining to the larger established parties, but also to marginal parties. These comparatively smaller parties are an important, albeit neglected, research object in their own right. They are of importance as, while they can only be appropriately studied using official electoral records, no one can tell whether or not they are indeed bound to remain marginal in the future.

Moreover, it would be desirable if election results at all levels of the political system were, as a rule, added to all regionalized data files provided by statistical offices. An even better alternative would be to set up a comprehensive database at the community (and city district) level, containing results of elections at all levels of the political system. A final desideratum concerns the data gained through the Representative Election Statistics program. Research possibilities could be substantially improved if these data would be made public not only at the level of the states, but also at the level of electoral districts. Participation studies, in their turn, could profit from access to process-produced data, such as data on extremist organizations collected by the State Offices for the Protection of the Constitution (*Landesämter für Verfassungsschutz*). Police records of demonstrations and estimated head counts of their participants could also greatly benefit participation studies. In the United States, such data have been successfully used to analyze the selection bias of mass media with regard to coverage of such protest events (McCarthy et al. 1996).

Survey data of high potential value for research into political attitudes and participatory orientations are constantly collected under the auspices of public agencies, such as the Press and Information Office of the Federal Government, and the public broadcasters ARD and ZDF. At present, only a small portion of these data are routinely submitted to the GESIS Data Archive. This seems hard to justify for data whose collection has been financed by public funds. Indeed, publicly funded projects can be seen as public property, which naturally the public has a right to access. In this regard, the German Freedom of Information Act (FOIA) clearly lacks bite. Under the US FOIA, data collected by public agencies are required to be made accessible to the public after three years at the latest. For three decades now, the Politbarometer surveys, as well as election studies conducted by the Forschungsgruppe Wahlen e.V. under the auspices of the ZDF, have been passed to GESIS. Cumulated over this long period of time, these data are a treasure trove for longitudinal political research, without which many important academic projects of electoral and participation research never would have seen the light of day. Starting with the data collected in 2008, the monthly Deutschland-Trend survey series conducted by Infratest dimap for the ARD is also made accessible through the GESIS Data Archive – a highly welcomed recent development. It would be desirable if this policy would also extend to other data collected under the auspices of the public broadcasters, in particular the exit polls conducted at elections, as already emphasized by Niedermayer (2001: 38). Moreover, in view of the increased interest in the role of media and communications for citizens' participation in politics, it would be highly desirable if the data collected by programs such as the ARD/ZDF's Mass Communication and Online Studies would be routinely

submitted to the public domain (on media data see also the advisory report by Meulemann and Hagenah in this publication).

Concerning the access of the scientific community to political surveys conducted by private survey institutes either for clients from the private sector or for their own purposes, one can only appeal for an increased readiness to submit these data to the GESIS Data Archive on the part of these institutes *and* their clients (who typically own the data). In this respect, at least one quite large recent project deserves highlighting: a public/private partnership between a group of academic researchers and the FORSA institute. Although it only partly improved data access for the scientific community at large, it provided a creative and original analysis of the dynamics of the 2002 parliamentary election by utilizing a very unusual and innovative database (Güllner et al. 2005). Private survey institutes also have been collecting data on media usage for decades. These data are of high interest for participation researchers, but so far the scientific community has only been granted limited access to these data (see the advisory report by Meulemann and Hagenah in this publication).

Turning to science-based programs of data collection, of the various ongoing replicative survey programs, two are of particular interest to researchers studying political participation and electoral behavior in Germany; namely, the German General Social Survey (ALLBUS, *Allgemeine Bevölkerungsumfrage der Sozialwissenschaften*) and the German Socio-Economic Panel (SOEP, *Sozio-oekonomisches Panel*). The ALLBUS is an indispensable resource for the long-term observation of trends in political participation and related topics. Fortunately, from its beginning, it has always carried political variables, and every ten years it has adopted political participation values and attitudes as core themes. It is strictly recommended to carry on with this rotating system in the future. For participation researchers in particular, it is considered vitally important to receive updates of key measures of political participation and related concepts at regular intervals (as well as data pertaining to new participatory phenomena). In doing so, the ALLBUS key working principle of combining replicative components with new, pre-tested instruments to catch up with recent societal developments seems highly appropriate.

In addition to the cyclical inclusion of political topics at a broader scale, each ALLBUS has always carried a small set of political indicators. However, the partial lack of long-term continuity with regard to these is disadvantageous. In the past, ALLBUS surveys included a number of important instruments, but several of them disappeared from time to time, either temporarily or permanently. Thinking about the future, in all upcoming waves of the ALLBUS, permanent inclusion of a small set of standard instruments would be extremely valuable for research into political participation and electoral behavior. Here, a commitment on the part of the ALLBUS program

to include them permanently as part of the essentials of the questionnaire would be welcome. These instruments should include:

- voting intentions and recall of vote decisions to record turnout and party vote at previous elections. These should pertain to national parliamentary elections. Additionally the recall question should ideally relate to the previous state and European elections;
- party identification (existence, strength, and party);
- party membership;
- left-right self-placement;
- interest in politics; and
- satisfaction with democracy.

It would be highly recommendable to include this same set of variables also into the standard SOEP questionnaire. This excellent database has so far not found many users among political scientists due to its glaring lack of political measures. Traditionally, the SOEP has carried only the standard indicator of party identification. It would therefore be highly welcome if the SOEP adopted at least the same small set of political standard instruments as essentials for its future surveys. For three reasons this would – at little cost – greatly enhance the utility of this impressive database: the uniqueness of the SOEP's panel design would open up unprecedented opportunities for analyzing change and stability of political orientations; the fact that it does not sample individuals, but households, would allow for analyses of the interdependence of individual orientations (Zuckerman et al. 2007); and, last but by no means least, its core content of socio-economic variables could be related to basic political attitudes (and their change).

While these steps towards increasing the value of the ALLBUS and SOEP programs would be highly desirable in view of the criterion of greatly enhanced synergy at little marginal cost, they could by no means replace a genuinely institutionalized program of research into citizens' political orientations. Although on the whole rather sanguine about the state of data provision and access for electoral and other political research in Germany, the 2001 KVI report emphasized a glaring gap in the otherwise very well developed German social science research infrastructure; namely, the lack of an institutionalized German National Election Study (GNES) that at each election reliably produces high-quality data as a public good (KVI 2001: 66; Niedermayer 2001: 33; Kaase and Klingemann 1994: 351-356; Kaase 2000: 32-34; Schmitt 2000; Gabriel and Keil 2005: 635-636). A significant step towards remedying this disadvantageous state of affairs has been made very recently. Starting with the 2009 Federal Election, a major research project, entitled the German Longitudinal Election Study (GLES), is funded by the German Research Foundation (DFG, *Deutsche Forschungsgemeinschaft*). The goal of this project is to cover the next three Federal Elections (Rattinger

et al. 2008). The GLES is exceptional in that it is a continuous program of empirical social research that meets the highest methodological standards, rests on a solid organizational base and transparent governance structure, enjoys the security of long-term funding, and is accountable. Furthermore, the GLES is open to the entire scientific community of academic empirical social researchers both with regard to the input side (i.e., with regard to developing the study design, questionnaires, etc.), and the output side (i.e., with regard to data availability and distribution). In light of this, it does not seem farfetched to claim that the GLES displays all the trademarks of the best election studies worldwide. In bearing with the previously described general trends, the GLES encompasses not only voter surveys, but also other components such as a candidate survey, interviews with party officials, and media content analyses. In this way, the GLES is in a unique position to place voting behavior in the broader context of the parties' campaign communications and the mass media's political coverage. Moreover, the GLES includes several longitudinal components (both repeated cross-sections and panels) that are to capture both the short-term dynamics taking place during election campaigns, and the long-term dynamics over entire electoral cycles. The study will also routinely include the CSES question modules (see above).

Overall, the GLES constitutes an important element of an emerging international infrastructure of high-quality data production and dissemination related to vitally important questions of the empirical foundations of democracy. It is conducted in close cooperation with the German Society for Electoral Studies<sup>5</sup> and GESIS. The former serves as an organizational network for linking the study to the scientific community, while the latter provides the study at all stages with logistic and methodological support, from developing research instruments to distributing the data via a web-based system. However, while being conducted according to the principles characteristic of high-quality National Election Studies worldwide, the GLES is still deficient with regard in one important respect – it will create unprecedented data infrastructure for the next three German national elections, but not beyond these. It would therefore be ideal if, in the long run, the study would be continued under the auspices of GESIS. In this respect, the GESIS could follow the model of the ALLBUS, which years ago mutated from a DFG project into an indispensable part of Germany's social science data infrastructure (within the remit of GESIS).

---

5 <http://www.dgfw.eu>

#### 4. Recommendations

- The ALLBUS is a replicative survey program of immense value to political research. For political scientists, it is essential that the ALLBUS carries on with its tried and tested rotating system of integrating broad political topics at regular intervals in the future. Moreover, it is strongly recommended that both the ALLBUS and the SOEP tag a small number of key political variables (listed above) as constant elements in their future question programs, ideally to be included in each wave. For the scientific organizations responsible for these two research programs, “value-adding” the ALLBUS and the SOEP in such a way would open the possibility of creating substantial amounts of synergy at little marginal cost.
- Responding to a grave deficit diagnosed by the 2001 KVI report, a determined collective attempt was recently initiated to close a glaring gap in the otherwise very well-developed German infrastructure of high-quality programs of replicative social science data collection. The ultimate goal of the effort leading to the German Longitudinal Election Study (GLES), which is at present funded by the DFG, is the institutionalization of a German National Election Study (GNES). In the long run, following the model of well-established continuous research programs, such as the ALLBUS and the SOEP (which are mostly designed to cater to the data requirements of sociologists and economists), this study should be granted permanent funding and become institutionally integrated into the over-arching infrastructure of the social sciences. Permanently establishing this study beyond the present DFG project, which covers the three German Federal Elections 2009, 2013, and 2017, would create an ideal supplement to the existing programs of replicative surveys in Germany. It also would generate unprecedented synergies with these pre-existing surveys. It is therefore to be recommended to policy-makers and research administrators to follow the model of other countries by providing the GNES with a stable financial basis of reliable public funding and an institutional embedding beyond the present GLES project. Ideally, this would occur by including it into the remit of GESIS.
- Concerning electoral data provided by the statistical offices, several improvements to data services are to be recommended. These include, for example, better provision of data on local elections, ideally as part of a comprehensive database at community (and city district) level. These data should contain results of elections at all levels of the political system. Further recommendations include detailed provision of electoral data on



marginal parties in computerized form, the addition of electoral data to regionalized data files, and publication of data from the Representative Election Statistics at the level of electoral districts. In addition, it is recommended to grant greater access to process-produced data pertaining to acts of collective (unconventional) participation.

- It is to be recommended that policy-makers establish a formal obligation for public agencies, including public broadcasters, to submit survey data collected under their auspices in due time to the public domain for purposes of secondary analysis. For this purpose, GESIS appears appropriate for archiving and disseminating such data. In particular, data of immediate relevance to participation, electoral, and political communication research is of importance.

## References:

- Dahl, R. (1972): *Polyarchy. Participation and Opposition*. New Haven.
- Falter, J.W. and Schoen, H. (Eds.) (2005): *Handbuch Wahlforschung*. Wiesbaden.
- Franklin, M. and Wlezien, Ch. (2002): *Reinventing Election Studies*. In: Franklin, M. and Wlezien, Ch. (Eds.): *The Future of Election Studies*. Amsterdam.
- Gabriel, O.W. and Keil, S. (2005): *Empirische Wahlforschung in Deutschland: Kritik und Entwicklungsperspektiven*. In: Falter, J.W. and Schoen, H. (Eds.): *Handbuch Wahlforschung*. Wiesbaden.
- Güllner, M./Dülmer, H./Klein, M./Ohr, D./Quandt, M./Rosar, U. and Klingemann, H.-D. (2005): *Die Bundestagswahl 2002. Eine Untersuchung im Zeichen hoher politischer Dynamik*. Wiesbaden.
- Kaase, M. (2000): *Entwicklung und Stand der Empirischen Wahlforschung in Deutschland*. In: Klein, M./Jagodzinski, W./Mochmann, E. and Ohr, D. (Eds.): *50 Jahre Empirische Wahlforschung in Deutschland*. Wiesbaden.
- Kaase, M. (2007): *Perspectives on Political Participation*. In: Dalton, R.J. and Klingemann, H.-D. (Eds.): *The Oxford Handbook of Political Behavior*. Oxford.
- Kaase, M. and Klingemann, H.-D. (1994): *Electoral research in the Federal Republic of Germany*. In: Thomassen, J. (Ed.): *The Intellectual History of Election Studies*, *European Journal of Political Research* 25 (3) (special issue), 343-366.
- Klingemann, Hans-Dieter (Ed.) (2009): *The Comparative Study of Electoral Systems*. Oxford.
- Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik (KVI) (Ed.) (2001): *Wege zu einer besseren informationellen Infrastruktur*. Baden-Baden.
- Lewis-Beck, M.S./Jacoby, W.G./Norpoth, H. and Weisberg, H.F. (2008): *The American Voter Revisited*. Ann Arbor.
- Lipset, S.M. and Rokkan, S. (1967): *Cleavage Structures, Party Systems, and Voter Alignments. An Introduction*. In: Lipset, S.M. and Rokkan, S. (Eds.): *Party Systems and Voter Alignments*. New York.
- McCarthy, J.D./McPhail, C. and Smith, J. (1996): *Images of Protest: Dimensions of Selection Bias in Media Coverage of Washington Demonstrations, 1982 and 1991*. *American Political Science Review* 61, 478-499.
- Milbrath, L.W. (1966): *Political Participation*. Chicago.
- Meulemann and Hagenah (2010): *Mass Media Research*. [In this publication].
- Mochmann, E./Oedegaard, I.C. and Mauer, R. (1998): *Inventory of National Election Studies in Europe 1945-1995*. Bergisch Gladbach.
- Niedermayer, O. (2001): *Wahlen und Parteien. Expertise für die KVI*. In: Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik (KVI) (Ed.): *Wege zu einer besseren informationellen Infrastruktur*. Baden-Baden.
- Pattie, Ch./Seyd, P. and Whiteley, P. (2004): *Citizenship in Britain: Values, Participation and Democracy*. Cambridge.
- Rattinger, H./Roßteutscher, S./Schmitt-Beck, R. and Weßels, B. (2008): *The Dynamics of Voting: A Long-term Study of Change and Stability in the German Electoral Process, Application to the Deutsche Forschungsgemeinschaft for long-term funding*. Bamberg/Frankfurt/Mannheim/Berlin.

- Romer, D./Kenski, K./Winneg, K./Adasiewicz, C. and Jamieson, K.H. (2006): Capturing Campaign Dynamics 2000 & 2004. The National Annenberg Election Survey, Philadelphia.
- Schmitt, H. (2000): Die Deutsche Nationale Wahlstudie – eher kollektive Aufgabe als aktuelle Realität. In: Klein, M./Jagodzinski, W./Mochmann, E. and Ohr, D. (Eds.): 50 Jahre Empirische Wahlforschung in Deutschland. Wiesbaden.
- Schmitt-Beck, R./Weick, S. and Christoph, B. (2006): Shaky attachments: Individual-level stability and change of partisanship among West German voters. *European Journal of Political Research* 45, 581-608.
- Thomassen, J. (Ed.) (1994): The Intellectual History of Election Studies. *European Journal of Political Research* 25 (3) (special issue).
- Thomassen, J. (Ed.) (2005): The European Voter. A Comparative Study of Modern Democracies. Oxford.
- Van Deth, J.W. (2003): Vergleichende politische Partizipationsforschung. In: Berg-Schlosser, D. and Müller-Rommel, F. (Eds.): Vergleichende Politikwissenschaft. Opladen.
- Van Deth, J.W./José Ramón Montero and Westholm, A. (Eds.) (2007): Citizenship and Involvement in European Democracies: A Comparative Analysis. London/New York.
- Zuckerman, A.S./Dasovic, J. and Fitzgerald, J. (2007): Partisan Families: The Social Logic of Bounded Partisanship in Germany and Britain. Cambridge.



## 6.2 Civil Society

*Mareike Alscher and Eckhard Priller*

Contact:

Mareike Alscher  
Eckhard Priller  
Wissenschaftszentrum Berlin für Sozialforschung  
Reichpietschufer 50, 10785 Berlin  
e-mail: alscher[at]wzb.eu  
priller[at]wzb.eu

## **Abstract**

Despite the obvious existence of civil society organizations (CSOs) and forms of civic engagement, the data available for this sector remain inadequate. This advisory report provides a comprehensive view of the current data situation, reveals existing gaps, and offers suggestions on how these gaps might be closed.

The empirical material currently provided by existing data sources – the Federal Statistical Office, statistics from the CSOs themselves, as well as special data and surveys – only register this area separately and to a limited extent, and even then not in a consistent manner. With respect to both CSOs and forms of individual engagement, the data situation is inadequate. Questions pertaining to whether CSOs will remain oriented toward civil society within a context of increased economic pressure cannot be answered. Similarly, it is nearly impossible to analyze whether civic engagement stands at odds to the increased tendency towards monetarization.

Since civil society will undoubtedly continue to gain in political and social importance, the long-term task will be to set up a meaningful and a predominantly self-contained system of data collection and provision. This goal can be reached by using the existing surveys and databases described in this paper.

Keywords: civil society; civil society organizations; civic engagement; Volunteers Survey; Johns Hopkins Comparative Nonprofit Sector Project; Handbook on Nonprofit Institutions in the System of National Accounts

## **1. Introduction**

The term “civil society” attracts a range of paraphrases and definitions. One of the most common definitions is action-oriented and focuses on four distinct attributes (Kocka 2003; Gosewinkel et al. 2004: 11). These attributes include (1) qualities of self-organization and independence; (2) an emphasis on actions taken in the public domain nurturing exchange, discourse, and understanding but also conflict; (3) the acknowledgement that conflicts and protests are included in this concept of civil society but they are associated with peaceful, non-violent, and non-military actions; and (4) a course of action that considers the common good above and beyond individual, specific, and particular interests.

This advisory report conceives of civil society in accordance with the logic of this field. As such, civil society can be perceived especially as characterized by (a) the self-organization of citizens and (b) their voluntary engagement in a number of organizational forms, such as clubs, associations, initiatives, or foundations. These organizations are generally regarded as the institutional core or infrastructure of civil society and are often collectively referred to as the “third” or “non-profit” sector as a way of separating them

from state and market sectors (Anheier et al. 2000). Civil society organizations (CSOs) thus constitute that area of society located between the boundaries of market, state, and family, and are characterized by their formal structures, organizational independence from state control, autonomous administration, non-profit approach, and voluntary engagement. Engagement in civil society organizations includes both unpaid voluntary work in traditional membership-based organizations and nonprofit-oriented activities in unconventional forms of organizations.

CSOs can be found in a variety of areas and perform diverse roles. Whether in recreational or cultural spheres, as part of social service facilities, or as other types of local, professional, and political advocacy groups (e.g., clubs, associations, foundations, not-for-profit Public Limited Companies, cooperatives, etc.), they have collectively become an essential part of society's workings.

As contemporary forms of civic self-organization and self-responsibility, CSOs possess considerable abilities in terms of the concentration, expression, and representation of interests. They are assigned responsibility for implementing important tasks, in promoting the development of democracy, providing welfare state services, as well as integrating citizens into coherent collectivities and thereby ensuring social cohesion.

A number of factors have led to the increased importance of this sector of society in recent years. On the one hand, citizens have become increasingly conscious of their own skills. On the other hand, social change has led to changes in social roles and functions, resulting in an increasingly stark division of tasks between state, market, and civil society. The growing significance of CSOs has manifested itself through increases in the number of CSOs, in the number of people working in them, and in the services they offer. At the same time, the number of voluntary workers also continues to grow.

Despite the obvious existence of CSOs and forms of civic engagement, the data available for this sector remains inadequate. Due to the relatively late development of the study of CSOs as an independent scientific discipline, the empirical information available on this constantly evolving sector is incomplete. Even official statistics and other data-providing information systems mark this area separately to a limited extent, and even then not in a consistent manner. For instance, CSOs and their services are often subsumed within the categories of state and economy, with data gathered from disparate surveys seldom taking their autonomous forms of organization into account.

CSOs tend to highlight the fact that they break down the classic dichotomy of state and citizen, replacing it instead with the three social spheres of state, market, and civil society. In the past, however, the autonomy of this sector did not prevent the use of CSOs for political ends in order to carry out those inconvenient tasks for which no one was – or considered themselves to



be – responsible. With this in mind, some social actors view CSOs as simply a form of cheap “repair service,” a way of balancing out the social deficits caused by the failure of the market, state, or family sectors.

Generally speaking, the growing demand for data on civil society can be explained by the increasingly autonomous significance of civil society in economic, social, and cultural life. Yet the current data situation is extremely complicated, not least because civil society has its own particular logic of action, and possesses unique functions and organizational structures, all of which have until now received only a modicum of direct attention and consideration. Data is lacking on the size of this sector, the extent of the services it offers, and its degree of socio-political integration. Current yet differentiated information is needed in order to more accurately define the significance of civil society, its development, and its contribution to providing solutions for current and future social challenges.

This expert report provides a comprehensive view of the current data situation, reveals existing gaps, and offers suggestions on how these gaps might be closed. Whereas in Germany relatively little data on civil society is available, other countries, such as the US, Australia, Italy, Belgium, and even Hungary, have progressed much further with regard to data collection and the long-term observation of civil society. Corresponding data is already an important component of these countries’ official statistics.

## **2. The current data situation in the civil society sector**

Empirical research on civil society can be divided into investigations aimed at three distinct levels. At the macro-level, CSOs are collectively analyzed as a field or sector. At the meso-level, research focuses on the CSOs, their specific tasks, and the way they function. Finally, at the micro-level, public activity in and for these organizations is investigated, with the key concerns in this context being membership, volunteering and donation behavior.

A considerable step toward improving the relatively awkward data situation was made with the Johns Hopkins Comparative Nonprofit Sector Project,<sup>1</sup> a large-scale, internationally comparative project with a scope spanning more than thirty countries. Under the coordination of the Johns Hopkins University Institute for Policy Studies (Baltimore, US), this project provided the results of data collected in Germany for the 1990 and 1995 reporting periods. The project was launched in 1990, and encompassed a

---

1 The project included formally structured, state-independent, and non-profit-oriented organizations. These organizations were also administered autonomously, funded to a certain extent by voluntary contributions, and could not in any sense be called an “administrative union” (Anheier et al. 1997: 15).

group of seven industrialized and five developing countries. In the meantime, the number of countries taking part has increased significantly. During the second phase (1995–1999), countries in North and South America, as well as both Western and Eastern Europe were heavily represented. Existing gaps in Africa and Asia have also been closed in more recent years thanks to the provision of additional country reports. Germany has been involved in the project from its inception.

The project collects quantitative data at national level on the structural dimensions of the non-profit sector, and investigates qualitatively how the sector is embedded within national structures. During the second phase, the German component of the project was located at the Social Science Research Center Berlin (WZB, *Wissenschaftszentrum Berlin für Sozialforschung*) and the Westphalia Wilhelm University of Münster's Institute for Political Science.

It was agreed that during the course of this international comparative project, empirical data on the CSOs would be collected according to the following targeted items:

- Number of CSOs
- Number of staff, based on number of hours worked (paid and voluntary staff)
- Financial volume
- Proportion of different funding sources within financial volume
- Expenditures
- Fields of activity
- Services provided

The well-established International Classification of Nonprofit Organizations (ICNPO) also formed part of the uniform research design. The nonprofit sector was then divided according to activity into twelve distinct fields, thereby allowing for an investigation into the internal structure of the sector. The ICNPO lists the following fields:

- Culture and recreation
- Education and research
- Philanthropic intermediaries and voluntarism promotion
- Health
- International
- Social services
- Business and professional associations, unions
- Environment
- Religion
- Development and housing
- Not elsewhere classified

The Johns Hopkins Project also developed a corresponding methodology, thereby establishing the essential groundwork for long-term observation. In collaboration with the Johns Hopkins University Center for Civil Society Studies, the United Nations Statistics Division produced the *Handbook on Nonprofit Institutions in the System of National Accounts*. This publication offers recommendations and guidelines for setting up national information systems. A host of countries (including Belgium, Italy, and France, among other European countries) have already adopted this approach. In Germany, however, no comparable administrative decisions and effective measures have been taken. Nevertheless, the implementation of this methodology is conceivable and could indeed be achieved on account of the close cooperation between the research community, the Federal Statistical Office, and CSOs.

In order to establish long-term and sustained observation of German civil society, data from official statistics as well as additional data stocks from CSOs, federal ministries, and other institutions and associations – including the research community – must be integrated. Despite endeavours to secure the continuous generation of reliable data on the social impact and performance of civil society in Germany, up to this point only partial and very basic data have been made available. Yet it would be possible to draw from sources ranging from official statistics, information from the CSOs, and, above all, data from scientific surveys. It is thus critical that the current data situation be fundamentally reshaped and improved; this must be set as a goal for the future. Greater coordination will be required in order to organize the amalgamation of the various data stocks. Moreover, scientific research, especially with regard to CSOs, must be undertaken. The current situation for the individual fields is as follows:

### *2.1 The Federal Statistical Office*

The Federal Statistical Office provides diverse statistics, although they do not fully conform to the methodology laid out in the *Handbook on Nonprofit Institutions in the System of National Accounts*. The information gathered from this source during the investigation referred to the number of CSOs, the number of staff, the CSO's financial volume, as well as the services offered, results, and capacities. These data do not, however, offer a full picture of the CSOs. Some of the surveys used to collect data are based on the 2003 German Classification of Economic Activities (WZ03) and use business entity classifications that are not consistent with the typical fields of activity and business entity classifications developed by the Johns Hopkins Project. Thus, the significance of these data is, generally speaking, limited. The following summarizes the individual data and data sources pertaining to CSOs within official statistics:

- a) *Economic accounts*
  - Data on gross value added and staff
- b) *Business register*
  - Data on turnover, number, and staff
- c) *Income tax statistics*
  - Data on financial volume (income) acquired through donations from private households
- d) *Corporation tax statistics*
  - Data on financial volume (income acquired through donations and the expenditure of donations; profit and loss information)
- e) *Survey on private schools (no current data available – last surveyed in 1995)*
  - Data on the number, income, and expenditure of private schools
- f) *University statistics (manual allocation of type of business entity necessary)*
  - Data on the number of universities, their staff, expenditure, income, and services provided
- g) *Research statistics of non-university research institutions*
  - Data on the number of institutions, their staff, expenditure, and income
- h) *Child and youth services statistics*
  - Data on the number of institutions, results, and income
- i) *Health service statistics*
  - Data on the number of institutions, their staff, services provided, and capacities
- j) *Continuous household budget surveys*
  - Data on financial volume (obtained through information on income, donations, and membership fees)
- k) *Income and consumer sample*
  - Data on financial volume (obtained through information on income, donations, and membership fees)
- l) *Time use survey (no current data available – last collected in 2002)*
  - Data on the engagement/volume of voluntary work

## 2.2 Statistics from Civil Society Organizations (CSOs)

Data received from umbrella organizations represent another important source of information for statistical analyses. However, the material provided from these sources is marked by certain gaps and irregularities. These gaps are caused by a number of factors. On the one hand, transparency is not particularly well developed in civil society organizations; the corresponding tax legislation means that only limited support is received from the state. On the other hand, the member organizations of these umbrella organizations – or even their regional branches at the level of the *Länder* – are themselves autonomous and independent legal entities and thus not obliged to provide data. Finally, incapability and noncompliance inevitably lead to gaps and loss of information.

At this point, it is useful to observe that when one considers the combined statistics available from all of the non-statutory welfare services in Germany (i.e., the voluntary welfare organizations of *Caritas*, *Diakonie*, the *German Red Cross*, *Paritätische*, and the *Central Welfare Office of Jews in Germany*), the combined statistical data in all the museums in Germany, and the database of the Association of German Foundations (*Bundesverband Deutscher Stiftungen*), it is very clear that CSOs could play a much larger role as potential suppliers of data in the future. The information gathered whilst conducting surveys such as the European Social Survey (ESS), which is a representative social survey that was first established in 2002/03 at the suggestion of the European Science Foundation (ESF), refers to the number of CSOs, the services offered and capacities, as well as the number of staff. The varying forms of data on CSOs in Germany are:

- a) *Overall statistics of the non-statutory welfare service sector*
  - Data on the number of institutions, staff, and their capacities
- b) *Overall statistical data for museums in Germany*
  - Data on the number of institutions, services provided, and number of staff
- c) *Association of German Foundations database*
  - Data on the number of foundations, their assets, and outputs

## 3.3 Special data and surveys focusing on a micro-level

### a) The Volunteers Survey

The German Volunteers Survey consists of a representative data collection in which around 15,000 German citizens over the age of 14 are queried about their level of civic engagement. To date, the survey has been carried out twice – in 1999 and 2004 respectively. The next survey is planned for 2009.

The survey's data – which have been scientifically verified – provides a number of opportunities for carrying out extensive analysis on the orientation, extent, and potential of civic engagement in Germany. At the same time, the survey provides information on the willingness of individuals to participate in civic activities (see Gensicke et al. 2006). Furthermore, the German Volunteers Survey offers insight into the motives behind civic engagement and the social structure of volunteerism. The differentiated data were collected according to socio-structural criteria.

b) The German Socio-Economic Panel (SOEP, *Sozio-oekonomisches Panel*)

By focusing on “social participation and time use,” the SOEP represents another crucial source of data in the combined statistical measurement of the level of engagement in civil society. Data from this source focuses on the types of engagement associated with certain forms of CSOs. At the same time, activities that fall within the purview of informal personal and community networks are also taken into account. Although the data gathered are not differentiated by specific fields of engagement, it is well-suited to the illustration of general trends and developments over time, and can also be used to implement time series analyses and analyses on socio-structural factors.

c) The IAB Establishment Panel

Data from the IAB Establishment Panel is primarily evaluated at the Institute for Employment Research (IAB, *Institut für Arbeitsmarkt- und Berufsforschung*). The survey gathers data from organizations that have at least one staff member subject to social insurance contributions. Therefore, the sample only contains those CSOs with paid staff. The survey thus provides information relating to staff and the CSOs. However, because it concentrates on economically active establishments, the broader spectrum of CSOs remains poorly represented.

d) European Social Survey (ESS)

As previously mentioned, the ESS is a representative social survey that was established at the suggestion of the European Science Foundation (ESF) and carried out for the first time in 2002/2003. In the first round, twenty-two countries participated (Austria, Belgium, the Czech Republic, Denmark, Finland, France, Germany, Great Britain, Greece, Hungary, Ireland, Israel, Italy, Luxembourg, the Netherlands, Norway, Poland, Portugal, Slovenia, Spain, Sweden, and Switzerland). With the exception of Switzerland and the Czech Republic, data on the level of engagement in civil society was pro-

vided for the remaining twenty countries. The long-term goal of the ESS is to investigate the interaction among political and economic institutions in transition, as well as the attitudes, convictions, and behavioral patterns of each country's population. The first round of the ESS focused on the themes, "Citizenship, Involvement, Democracy." The 2002/2003 survey used a four-step approach to gathering data on civic engagement, including: (1) being a member of a CSO, (2) working for a CSO, (3) donation behavior, and (4) civic engagement within a CSO.

### **3. Gaps, progress, developments, and tendencies of the current data situation**

This portrayal of the current data situation makes it abundantly clear that a comprehensive and developed information system on civil society simply does not exist at this point. While individual engagement can be analyzed through different scientific surveys, other areas show distinct deficits. Data gaps exist particularly where it concerns CSOs and their concrete fields of activity. To date, the current picture – including over 600,000 associations, more than 14,000 foundations, around 8,000 registered cooperatives, and numerous other organizations – is more than a little hazy. Information on newly established or disbanded CSOs can, as a rule, only be found by searching through special registers existing for different forms of organizations. One particularly significant gap is the absence of broader scientific surveys that cover all of the different organizational forms of CSOs; other countries (e.g., Austria) have already embraced this approach in recent years and integrated such broad scientific surveys on CSOs as part of their federal statistics. These extensive surveys offer insight into the dynamic changes in the orientation and activities of different CSOs. In light of the increasing economic pressure on organizations, which tend to react to such strain by improving their management techniques or by tapping into additional financial resources (e.g., donations), the importance of extensive surveys offering insight into the inter-temporal changes of CSOs has grown. This increased financial strain simultaneously raises the question of whether CSOs intend to remain oriented toward civil society. A change in orientation could lead CSOs to regard the engagement of the population in civil society as less important. Economic factors can therefore lead to the neglect of civic engagement and volunteerism on the part of the CSOs.

Moreover, irrespective of all that might be done at the organizational level (of CSOs), many questions that concern the civic engagement sector at the individual level remain unanswered: either no data are available or existing data cannot come up with adequate answers. As a consequence, it is

almost impossible to analyze whether civic engagement stands at odds with an increased tendency towards monetarization. This would seem to confirm theories which declare an increased tendency towards the dissolution of boundaries between paid and unpaid work vis-à-vis gainful activity. Research is also needed to ascertain whether value change is taking place within the context of civic engagement and whether forms of a stronger, non-organized engagement, one which requires no concrete membership, are becoming increasingly prevalent.

Alongside these obviously significant gaps in the data, however, some developments and tendencies can be detected that point to improved data collection and analysis.

#### a) Civil Society Data Collection Project

Due to the grossly inadequate data situation that exists with regard to civil society, several foundations have decided to sponsor a new intervention. The aim of this project, which will run until 2010, is to establish a reporting system based predominantly on the Federal Statistical Office's data stocks. The reports will focus on providing an economic balance sheet and social service profile for CSOs. The project is located at the *Stifterverband Wissenschaftsstatistik GmbH*, the research and development branch of the Donors' Association for Promotion of Science and the Humanities in Germany (*Stifterverband für die Deutsche Wissenschaft*), and will provide basic data for further investigations. Based on the concepts and methods of the Johns Hopkins Comparative Nonprofit Sector Project, with existing international standards that have developed in the meantime, the midterm goal of the project is to establish a national accounting satellite system.

#### b) Report on Donation Behavior

So far there have been a number of investigations that deal with donation behavior. Among these we find the "Donations Survey" (*Spendenmonitor*) by EMNID and the "GfK Charity\*Scope" survey of the GfK Group, an international market research company.

Amidst calls for greater transparency in the donation sector and increasing competition among non-profit organizations, plans are underway to publish a national report on donation behavior. The Social Science Research Center Berlin (WZB, *Wissenschaftszentrum Berlin für Sozialforschung*) has developed the report methodology which contains, among others, information on donation volume, donors, purposes, and motives. The project is to be conducted by the German Central Institute for Social Issues (DZI, *Deutsches Zentralinstitut für soziale Fragen*).



#### **4. Future requirements and perspectives for civil society data**

Civil society will undoubtedly continue to gain in political and social importance as we move into the future. It is therefore highly likely that the demand for data and analyses will also increase. The long-term task, as it has been in many other areas of society, will be to set up a meaningful and predominantly self-contained data collection and provision system. This goal can be reached by using the existing surveys and databases that have been described in this paper. Along with more substantial and better methods of coordination, the criteria and categories for civil society must be integrated into other data collection activities. Considerable progress could be made by ensuring that the type of business entity represented by CSOs, or their nonprofit orientation, is considered as a specific criterion throughout. Subsequent analyses could also be strengthened by integrating questions about civic engagement into other specific large-scale surveys (e.g., the annual Microcensus). Experience in Austria has shown that using this approach significantly improves the availability and quality of data.

More effort must be directed toward carrying out larger surveys on CSOs. The impact of research in this field and the evaluation of particular structures and practices will have increasing significance.

#### References:

- Anheier, H.K./Priller, E. and Zimmer, A. (2000): Die zivilgesellschaftliche Dimension des Dritten Sektors. In: Klingemann, H.-D. and Neidhardt, F. (Eds.): Zur Zukunft der Demokratie. Herausforderungen im Zeitalter der Globalisierung. WZB-Jahrbuch 2000.
- Anheier, H.K./Priller, E./Seibel, W. and Zimmer, A. (Eds.) (1997): Der Dritte Sektor in Deutschland. Organisationen zwischen Staat und Markt im gesellschaftlichen Wandel. Berlin.
- Gensicke, Th./Picot, S. and Geiss, S. (2006): Freiwilliges Engagement in Deutschland 1999 - 2004. Wiesbaden.
- Gosewinkel, D./Rucht, D./van den Daele, W. and Kocka, J. (2004): Einleitung: Zivilgesellschaft – national und transnational. In: Gosewinkel, D./Rucht, D./van den Daele, W. and Kocka, J. (Eds.): Zivilgesellschaft – national und transnational. WZB-Jahrbuch 2003.
- Kocka, J. (2003): Zivilgesellschaft in historischer Perspektive. In: Forschungsjournal Neue Soziale Bewegungen 16 (2), 29-37.
- van Deth, J.W. /Montero, J.R. and Westholm, A. (Eds.) (2007): Citizenship and Involvement in European Democracies. A comparative analysis. London and New York.

## 6.3 Culture

*Jörg Rössel and Gunnar Otte*

Contact:

Jörg Rössel  
Gunnar Otte  
Universität Zürich  
Soziologisches Institut  
Andreasstrasse 15  
8050 Zürich  
Switzerland  
e-mail: [roessel\[at\]soziologie.uzh.ch](mailto:roessel[at]soziologie.uzh.ch)  
[otte\[at\]soziologie.uzh.ch](mailto:otte[at]soziologie.uzh.ch)

## Abstract

The term “culture” is notorious for its multitude of meanings. This advisory report strictly focuses on culture in terms of the arts. We adopt a sociological as well as an economic perspective. Research questions are subdivided into three spheres: artistic production and its organization; the distribution and valuation of culture; and the consumption and reception of culture. The data requirements and the availability of adequate data vary substantively, depending on artistic branches (music, performing arts, etc.) and specific research questions. In order to make the empirical investigation of culture a flourishing field, we recommend the following improvements in data infrastructure: first, comprehensive surveys of artists on the one hand, and cultural consumption on the other, should be carried out with the support of public funding; second, a national cultural statistic should be established, illuminating the size, impact, and evolution of the cultural sector in comparative perspective; third, the public availability of organization-level data as well as communal surveys on cultural production and consumption issues should be improved; fourth, the transparency of existing data sources and their accessibility should be improved by archiving them centrally, e.g., at the GESIS<sup>1</sup> Data Archive.

Keywords: culture, arts, artists, production, distribution, consumption, reception, cultural sector, cultural industries

## 1. Definition of culture

The term “culture” is notorious for a multitude of definitions. In our advisory report we strictly focus on culture in terms of the arts. Issues that are sometimes included under the superordinate concept of culture, like religion, ideologies, values, norms, and patterns of everyday life, are not considered in this paper. Including these topics would necessarily lead to a superficial treatment of each because of the numerous and heterogeneous data sources in these areas. However, even the concept of “arts” itself must be differentiated. Generally speaking, the arts include objects and services of primarily aesthetic expression. These are, first, differentiated according to the implied aesthetic criteria. In public as well as scientific discourse, high culture, popular culture, folk culture, and youth culture are typically distinguished even if these terms are difficult to mark off in their boundaries (Gans 1974; Schulze 1992; Hügel 2003). While folk, popular, and youth culture are often normatively devalued, all of these aesthetic forms have to be included in empirical research from a value-free scientific point of view. This is because conceptions of beauty are socially constructed and historically variable.

---

1    Leibniz Institute for the Social Sciences (*Leibniz-Institut für Sozialwissenschaften*).

Secondly, the arts have to be differentiated into core branches like music, performing arts, literature, visual arts, and film. Since these areas exhibit varying forms of social organization (Deutscher Bundestag 2007, chap. 3), most research questions have to be applied separately to these fields. These internal differentiations of the arts lead to a multiplication of the data sources required for empirical research.

## **2. Theoretical developments and research questions**

In the last major German publication of the sociology of the arts, Gerhards (1997: 7) concluded that this field is not at all established in German sociology. The situation has remained nearly unchanged. German sociology has not participated in the international boom of the sociology of the arts and culture.<sup>2</sup> Most sociologists in the field prefer qualitative methods – if they do empirical research at all. To be sure, qualitative research and case studies are important complements of the standardized data that we focus on here. We do not further discuss this strand of research because it almost always involves primary data collection. Due to this basic research orientation and because of other reasons to be described in section 3, adequate data enabling scholars to tackle central research questions are scarce.

Contemporary sociology of the arts and culture is not about the interpretation of artistic content. Although this orientation can still be found in the literature, the main focus is – in accordance with Max Weber – on the description, understanding, and explanation of social action related to goods and services of primarily aesthetic expression. Research questions are usually subdivided into three different spheres of action: first, artistic production and its organization; second, the distribution and valuation of culture; and third, the consumption and reception of culture (Becker 1982; Blau 1988; Gerhards 1997; Schneider 1993; Zolberg 1990). Apart from sociology, the field of cultural economics has developed recently. Therefore, we include research questions and data requirements of economists of the arts and culture in our report (Blaug 2001; Caves 2000; Frey 2000; Ginsburgh and Throsby 2006; Throsby 1994).<sup>3</sup>

---

2 In the most important journal of empirical research in the arts (Poetics), we find only one article from Germany and one from Austria in the issues from 2003 to 2007. In comparison, economists from Germany and Austria have published nine articles in the major journal in the field of cultural economics (Journal of Cultural Economics).

3 Although being very important for the explanation of phenomena related to the arts and culture, we do not discuss psychological research because it is mainly based on experimental data.

## *2.1 Artists and production of culture*

The sociology and economics of artistic production deal with four broad research questions. They are, first, concerned with the socialization, recruitment, and training of artists, as well as the social inequalities connected with artistic career paths that vary in terms of social class background, general and artistic education, gender, ethnicity, earnings, and social security (Menger 1999; Caves 2000, chap. 4; Janssen 2001; Haak 2008). Second, inter- and intra-individual variations in living and working conditions are supposed to have an impact on artistic output, creativity, and aesthetic development. Both questions necessitate longitudinal data that link artists' labor market positions and integration in artistic networks with their aesthetic expression and that track stability and change over their life courses (Thurn 1983; Simonton 1997; Bourdieu 1999). Third, scholars are interested in the institutional organization of artistic production, its conditions, and consequences. They try to explain why organizational forms of artistic production vary enormously between cultural branches and between countries. They also try to assess the impact of these variations on artistic outcomes: bureaucratic organization vs. short-term projects, public vs. private funding, types of contracts between artists and support personnel, organizational structures dealing with market uncertainty (Peterson 1976; Caves 2000; Dowd 2004; Deutscher Bundestag 2007, chap. 3; Gebesmair 2008, chap. 4). Finally, the production of culture may be considered from a macro perspective. The importance of culture for the economy has become an important issue for research and official statistics as several German states and cities have published reports on the cultural sector (*Kulturwirtschaftsberichte*). Currently, a lively political debate centers on the establishment of a unified statistic of the cultural sector in Germany and Europe (Statistisches Bundesamt 2004; Deutscher Bundestag 2007, chap. 5; Eurostat 2007, part II).

## *2.2 Distribution and valuation of culture*

Producing a good or providing a service does not make it art. The status of art is based on the authentication of a good or service as art by accepted authorities like critics, curators, gallery owners, and ministries of culture. Therefore, the development of aesthetic criteria to evaluate art and the canonization of artists and art forms is a central research area (Bever 2005; Baumann 2007). However, cultural authorities do not only consecrate goods or services as art; they recommend and interpret art works for the lay public and are thereby actively engaged in the creation of markets for art and in price formation on these markets (Shrum 1997; Caves 2000, chap. 12; Beckert and Rössel 2004; King 2007). Social scientists depend on data about

cultural authorities and critical evaluations, which are essential for artists' reputations, as well as on market data, like prices for art works and box office results, which reflect their commercial successes.

Other actors and organizations, like gallery owners, museums, concert halls, and radio stations, are decisive for the supply and distribution of cultural goods and services. They perform gate keeping functions in artistic fields, create artistic repertoires, and thereby advance or hamper artistic careers (Greenfeld 1988; Mark 1998; Giuffre 1999). Again, we find a vast array of different organizational forms in the distribution and valuation of culture. Explaining why certain forms emerge and which consequences they imply are central topics for both sociology and economics (Frey 2000; Kirchberg 2005). Data on cost and finance structures of institutions are of further importance from an economic perspective, as they enable researchers to evaluate the efficiency of the provision of culture, e.g., theatres in the profit- vs. non-profit sector.

### *2.3 Consumption and reception of culture*

Questions of the consumption and reception of culture have generated the bulk of empirical studies in sociology. A lot of research has been devoted to inequalities of social class, gender, ethnicity, age, and generation in cultural consumption, especially with regard to the use of publicly funded cultural institutions (Dollase et al. 1986; Klein 1990; Rössel et al. 2005; Kirchberg 2005; Bourdieu et al. 2006). However, much of this research is of a rather descriptive kind and the data usually collected do not allow scholars to test rival hypotheses and reveal explanatory mechanisms. For example, there is a long-standing and still open debate about whether the well-known educational effects on high-culture consumption are based on information-processing or status-seeking mechanisms (Ganzeboom 1982; Otte 2008). In order to fill these research gaps, scholars are dependent on adequate survey data containing theoretically derived indicators. In particular, longitudinal individual-level data are of prime importance for the analysis of the biographical formation of aesthetic preferences (Hartmann 1999; Katz-Gerro et al. 2007). In this respect, findings in the sociology of culture are of a more general interest, as the origin of preferences constitutes a central question in the behavioral sciences. Closely related is research on the symbolic boundaries people draw in order to express their likes and dislikes for different aesthetic forms and genres (Lamont and Molnár 2002). A major international debate circles around the thesis of so-called “omnivorous” tastes. This implies a reorganization of traditional taste hierarchies: the educated classes in Western societies are said to have stopped using high culture as the main aesthetic format of distinction vis-à-vis the lower classes, and instead to have broadened their taste repertoire with popular genres and to display wide-



ranging competences as new status-markers (Peterson 2005). High-quality time series data are needed to study such preference and consumption patterns over time, comparative data are required to find out about international variation.

### 3. Databases and access

In comparison with other research areas, the data infrastructure in the field of culture is not well-institutionalized. In academia, there has been no establishment of a research program based on comprehensive, recurrent nation-wide surveys on cultural production and consumption, let alone panel studies. In official statistics, the cultural domain falls under the sovereignty of the federal states and communes. A standardized, unified cultural statistic on the national level is nonexistent. In this regard, the conclusion of the 2001 KVI report still holds: reporting on cultural issues is rather unsystematic (Weishaupt and Fickermann 2001: 50).

This does not mean that there is a scarcity of data on culture. Rather, as has been noted by the KVI report (2001: 16) for other fields, the current situation resembles a fragmented mosaic of various data lacking comparability, being frequently intransparent or inaccessible, and thus inhibiting cumulative research efforts. We will shed light on this situation according to the three main spheres of research that we distinguished in the last section. We consider both aggregate- and individual-level data from various sources.<sup>4</sup> Although we wish to emphasize the much greater analytic potential of individual-level data for most research questions, aggregate-level data are valuable especially for some applied and policy-relevant questions.

#### 3.1 Artists and production of culture

Artists' socialization processes and careers are of interest from a double perspective: the formation of aesthetic expression over the life course and social as well as material inequality within the cultural field. Both questions can be addressed most systematically by using surveys tracking artists' retrospective careers and using a research design like the German Life-History Study (Mayer 2008). Assessing individual artists' development of aesthetic expression and productivity can be further improved by linking respondent data to documentary sources on art works for a subset of cases. While, to our

---

4 Aggregate data are data that have been aggregated from smaller units of analysis and cannot be easily disaggregated again. Individual-level data, in our case, refer to both persons and organizations.

knowledge, such datasets are nonexistent, the situation is somewhat better for questions of inequality. In order to study patterns of intergenerational social mobility and reproduction among artists, cumulative ALLBUS- and SOEP-data may, in principal, be used (Jonsson et al. 2007). However, the number of respondents is very small; artistic branches cannot be differentiated. The German Microcensus has the great advantage of large numbers, but lacks sufficient biographical information. Still, it has been used to investigate the effects of socio-demographic variables on artists' employment relationships and earnings (Haak 2008).

In this context, limitations become apparent regarding data from official statistics of artists' earnings and material living conditions.<sup>5</sup> The main data sources are the German Microcensus, the Employment Sample of the Institute for Employment Research (IAB, *Institut für Arbeitsmarkt- und Berufsforschung*), the statistics of the German Artists' Social Insurance Company (KSK, *Künstlersozialkasse*) and sales tax statistics (*Umsatzsteuerstatistik*). They differ substantially in their coverage of the artist population. The Microcensus, for example, counts everyone who works at least one hour per week in his or her main occupation as employed – and thus includes individuals regarded as not employed by the IAB's Employment Sample. The latter considers employees who are subject to social insurance contributions, work at least 15 hours per week, and earn at least € 400 per month. Because it does not cover, among others, the self-employed, it may be combined with KSK statistics, a social insurance institution open (on a voluntary basis) for self-employed artists earning at least € 3,900 annual artistic income. The sales tax figures include businesses with more than € 17,500 annual turnover and thus exclude self-employed “starving” artists. The databases also differ in their classifications of cultural occupations: the Microcensus defines occupational affiliation according to respondents' self-assessments, the Employment Sample according to employers' reports, and the sales tax statistics according to tax inspectors' assignments. The Employment Sample and sales tax statistics can be broken down to low occupational levels, but they do not contain enough individual-level information to model explanatory variables in statistical analyses. The Microcensus as a household survey may be preferable in this respect, but it does not offer a fine grouping of occupations. None of these data sources properly comes to grips with multiple job holdings and the mixture of dependent and self-employment typical for the artist population (Haak 2008).

---

5 Haak (2008, chap. 3) gives a detailed discussion of the problems the Microcensus and the IAB's Employment Sample have. Apart from problems due to the incomplete coverage of the artist population and the aggregation of occupational subcategories, inconsistencies of educational variables, censored income variables, the lacking differentiation of income sources, and multiple job holdings are considered problematic.

Because of these coverage, classification, and measurement problems, estimates of the number of artists, their education, and earnings differ depending on the data used (Haak 2008, chap. 4; Deutscher Bundestag 2007: 289ff).<sup>6</sup> Against this background, an explicitly designed survey on the living conditions of artists would be highly desirable. More than thirty years after the pioneering work of Fohrbeck and Wiesand – “Autoren-” and “Künstler-report” (1972; 1975) – primary data still need to be collected on a large representative sample of artists and other persons close to the creative core of the cultural sector.<sup>7</sup> Nevertheless, official statistics will be important for continuous social reporting and construction of time series. Therefore, an integration and standardization of current statistics is needed.

A similar conclusion holds for the effects of the cultural sector on the economy, usually measured in turnover and employment figures. Problems of definition, classification, and comparability, pervade the currently popular *Kulturwirtschaftsberichte* (Weckerle et al. 2003; Statistisches Bundesamt 2004; Deutscher Bundestag 2007, chap. 5). The relevant target population extends far beyond those occupations that would count as “cultural” according to our definition. Usually, all self-employed and dependently employed people in the production and distribution of goods and services in the visual and performing arts, publishing, press, radio, television, music, film, architecture and design, cultural education, and maintenance of cultural heritage, are subsumed under the label “cultural industries.” This already broad category is sometimes expanded to include those in the advertisement, software, and games industries, and correspondingly entitled “creative industries.”<sup>8</sup> There is disagreement, however, on the following (Deutscher Bundestag 2007: 340ff): is cultural employment in the public sector to be counted among the cultural industries? Are non-profit, voluntary, and lay cultural

---

6 The boundary problem of who is an artist is difficult to solve because the arts are not as professionalized as other occupations (Karttunen 1998). A minimum proportion of income earned or of hours worked can serve as criteria. In addition, the subjective self-categorization as an artist, educational credentials, and institutional affiliations, have some plausibility. Notably, artistic status is professionally or publicly ascribed and undergoes historical change. Current examples of boundary cases – sometimes legal cases about inclusion in the KSK – comprise assistant directors, disc-jockeys, web designers, and curators. A classic, prevailing controversy is related to the boundary between arts and crafts (Becker 1982: chap. 9).

7 In connection with an inquiry into “Culture in Germany” (Kultur in Deutschland), a large-scale online and mail survey addressing self-employed artists was launched by a culturally committed consultant, Christian Scheibler. In various aspects, e.g., sampling procedure and questionnaire construction, it did not follow standards of scientific research (Kressin 2008). This example highlights the urgency of a methodologically sound “status-of-the-artist” survey in Germany. Otherwise, we see the danger that the artist population, known to be particularly hesitant to provide personal information, may lose trust in future survey efforts.

8 Söndermann (2005) combines data from the Microcensus, Employment Sample, and sales tax statistics, to make estimates of employment in the cultural industries, differentiated by branches, employment status, regional distribution, and development from 1999 to 2004.

activities to be included (e.g., choirs, music clubs, theatre groups) – and how can they be reliably captured? Are both a narrow and a broad definition necessary, and if so, which cultural branches belong to the core of the cultural sector? Are whole branches to be incorporated or just the creative parts of them (e.g., writers, but not printers)? A consensus on these questions is needed to guarantee the comparability of future reports on cultural industries in different countries, federal states, and cities.

The Federal Statistical Office (Statistisches Bundesamt 1994; 2004) has suggested a conception for a nationally unified culture statistic and illustrated the potential of standardized indicators in a recent publication (Statistische Ämter des Bundes und der Länder 2008).<sup>9</sup> Cultural statistics have also been presented at the European level (Eurostat 2007). They are based, among others, on the EU Labour Force Survey, Structural Business Statistics Survey, EU Household Budget Survey, Harmonized European Time Use Study, and the Eurobarometer, but provide a rather incomplete and tentative picture. From a scientific point of view, cultural statistics and reports on the economy of culture encompass important macro indicators, allowing researchers to make spatial-temporal comparisons and to identify broad trends. The more aggregated the data are, however, the less potential they have for revealing social processes at the micro-level within the cultural sector.

### *3.2 Distribution and valuation of culture*

Research on the distribution and valuation of culture requires organizational and archival data. In order to learn more about the types of cultural products and services which are distributed, longitudinal data on artistic repertoires of institutions and companies are needed, e.g., repertoires of theatres and orchestras, inventories and exhibitions of museums, circulation and sales figures of books and records. In the case of public sector institutions, especially theatres, operas and orchestras, such information is accessible via archival documentation of single institutions and increasingly via internet websites. This information can then be used to generate datasets (Mark 1998; Gerhards 2008). The situation becomes worse, however, the smaller the organizations are (e.g., free theatres) and the more profit-oriented they are (e.g., musicals, record companies).

For such purposes, publications of professional associations are important sources. These include, for example, the Institute for Museum Research

---

9 For the Federal Statistical Office's enquiry into "Culture in Germany" (Statistisches Bundesamt 2004: 208-311), it produced an advisory report, which contains an extensive account of the official data sources currently available for the creation of a federal culture statistic.

(*Institut für Museumsforschung*), the German Theatre and Orchestra Association (*Deutscher Bühnenverein*), the German Publishers and Booksellers Association (*Börsenverein des Deutschen Buchhandels*), the Confederation of the Film Industry (*Spitzenorganisation der Filmwirtschaft*), the Association of the German Music Industry (*Bundesverband Musikindustrie*), the German Confederation of Socio-cultural Centers (*Bundesvereinigung sozial-kultureller Zentren*), and the German Choral Association (*Deutscher Chorverband*). Data reported in the annual reports of such institutions are based on (a) routine surveys of samples of cultural institutions or of their member organizations, (b) questionnaires on special topics, and (c) sales tax statistics. Official statistics often rely on these figures in their yearbooks. Additionally, collecting societies, such as GEMA, GVL, and VG Wort,<sup>10</sup> hold data on musical and literary publications.

A central shortcoming of these data sources is that they are subject to high aggregation levels and information scarcity. Reports usually aggregate figures of single organizations on turnover, ticket prices and sales, utilized seat capacity, persons employed, and other indicators, without differentiating sufficiently between organizational forms and sizes. For scientific purposes, disaggregated organizational-level data are most desirable because they allow researchers to classify organizations according to the question at hand. Also, information about concrete repertoires is frequently missing. If concrete products are mentioned, they are often confined to successes, e.g., the top 50 movies of the year. However, similar annual “flop” lists (in combination with production costs) would be of equal importance because they constitute negative cases for comparative analyses.

For economic analyses, more data on organizational cost and finance structures are of high importance. Most detailed information can be found in the theater statistics of the German Theatre and Orchestra Association (*Deutscher Bühnenverein* 2008a; 2008b). Down to the organizational level, it provides data on repertoire, performances, seat capacity, visitors, personnel, revenue, and cost structures, as well as prices. This detailed data provision could serve as a model for the museum statistic (*Institut für Museumsforschung* 2007). A further improvement would be electronic access to these organizational-level data because, otherwise, data preparation for statistical analyses is very cumbersome.

A second problem has to do with organizational coverage, sampling, and response bias. The coverage of cultural institutions and organizations is often

---

10 GEMA (Gesellschaft für musikalische Aufführungs- und mechanische Vervielfältigungsrechte) is a non-profit organization representing the copyrights composers, lyricists, and music publishers, GVL (Gesellschaft zur Verwertung von Leistungsschutzrechten) is an association representing the copyright interests of performing artists and record manufacturers, and VG Wort (Verwertungsgesellschaft Wort) is a copyright association of authors and publishers.

intransparent or – as in the case of the theater statistic – biased towards large, professional, publicly funded or member organizations. Precise methodological information on the target population of organizations, sampling issues, and the data collection methods of participating organizations are necessary to assess the quality and information content of the data.<sup>11</sup> A potential problem of unclear incidence might result from organizational interests of professional associations and their influence on questionnaire content, question design, organizational population covered, and statistical reporting. The availability of data from professional associations varies between cultural branches. There is, for example, relatively rich information for theatres, museums, the phonographic and film industries, some information for publishing, and poor information for socio-culture and the primary market for visual arts (Statistisches Bundesamt 2004: 312-423).

Data on valuation processes in the arts are most useful when collected through content analyses of documentary sources. These include reviews in journals, newspapers, and art history books, as well as coverage and accounts of artistic products in school books (Bever 2005). Scientific access to these sources exists via the German National Library (*Deutsche Nationalbibliothek*), other libraries, and archives of journals and newspapers. It seems important to broaden the coverage of libraries and archives to smaller art periodicals.

### 3.3 Consumption and reception of culture

As mentioned above, most empirical studies in the sociology of culture focus on consumption and reception issues and utilize survey data. Modules on culture in our sense appear in various surveys and are largely accessible via the GESIS Data Archive. These typically include general social surveys like the Welfare Survey (*Wohlfahrtssurvey*) 1993 and the German General Social Survey (ALLBUS, *Allgemeine Bevölkerungsumfrage*) 1998, youth surveys like the Shell-Jugendstudie, comparative surveys like the Eurobarometer 67.1/2007, studies on media consumption like “Massenkommunikation I-VI,” and surveys on reading conducted by the Stiftung Lesen in 1992, 2000, and 2008 (not available at the GESIS Data Archive). These studies usually ask respondents about the frequency of consumption of a set of artistic goods and services. However, they do not go into details of the specific contents being consumed and the ways they are consumed, while these studies sometimes

---

11 The advisory report of the Federal Statistical Office (Statistisches Bundesamt 2004: 312-423) describes non-official data sources extensively. Among the statistics of professional associations, the museum statistic is also exemplary in its methodological documentation and its efforts to achieve a high response rate. Taking into account that methods of annual visitor counts vary enormously between museums – from cash registers to pure estimates – however, reliability problems even in quite simple indicators become apparent.

employ multidimensional categories like “theatres and concerts” and they contain little information on the biographical formation of consumption and reception practices and their embeddedness within social networks. Thus, currently available data are not suitable to test rival hypotheses about the origin and development of aesthetic preferences. Nor are they sufficient to reconstruct modes of cultural education or the ways symbolic boundaries are drawn. It is therefore clear that academic research in Germany has hitherto not developed a comprehensive, recurrent survey on the cultural consumption and reception of the general population.<sup>12</sup>

The survey that comes closest to an institutionalized reporting on cultural tastes and activities of the population is the “Kulturbarometer.” This survey has been conducted eight times since 1991 by the Centre for Cultural Research (ZfKf, *Zentrum für Kulturforschung*), Bonn – the same institute that was responsible for the “Künstlerreport.” Although the published results of these surveys are sometimes accompanied by extensive and informative tables, the data are currently not accessible for secondary analyses. Because the ZfKf is built on project-specific funding, continuous cultural reporting is currently not ensured.<sup>13</sup>

Since the 1990s, survey research on local-level cultural participation has flourished in cities and municipalities. These data are usually compiled by local statistics agencies or, sometimes, by academic or commercial research institutes on behalf of local authorities. The existing data infrastructure is very intransparent because these research activities are scattered all over the country, results are not made accessible to the wider public, and data are not centrally archived. There are efforts by the Union of German Municipal Statisticians (VDSt, *Verband Deutscher Städtestatistiker*) to coordinate and standardize surveys in order to achieve better comparability of local results via programs such as KOSIS (“Kommunales Statistisches Infrastruktursystem”) and UrbanAudit. Recommendations for questionnaire construction have also been made (Deutscher Städtetag 1994). Notably, a database for research reports and questionnaires of communal surveys (“komm.DEMOS”) is located at the German Institute of Urban Affairs (Difu, *Deutsches Institut für Urbanistik*), Berlin (Bretschneider and Schumacher 1996). This database

---

12 The situation is, as far as we know, not much better in other countries. In the US, the replicative survey SPPA (“Survey on the public participation in the Arts”) was conducted in 1982, 1992, and 2002 enabling scholars to make temporal comparisons (DiMaggio/Mukhtar 2004). However, it was an add-on to other surveys and impaired by methodological problems (Peterson 2005). Quite extensive surveys on culture are carried out in the Netherlands, but we do not know about recurrent social reporting on this topic.

13 As a response to our request for opening its databases for scientific secondary analyses, the director of the ZfKf, Andreas Johannes Wiesand, signaled a general willingness to make primarily older data available to the GESIS Data Archive. However, some of them – e.g., data of the “Künstlerreport” – frequently do not exist in electronically readable form. Resources are needed to convert them.

currently comprises about 2,000 standardized study descriptions, 400 of which are culture-related. It is accessible free of charge primarily to communes having provided financial contributions (“Zuwenderstädte”), but not for the scientific community. Komm.DEMOS, however, does not archive the survey data itself, nor does any other central archive for communal surveys exist.<sup>14</sup> We expect communal survey data to vary in quality, depending on issues of survey administration and methodological rigor. Individual-level data of well-organized surveys are of great scientific value as they are more context-sensitive than nationwide surveys. They entail information on a broad range of the locally available cultural infrastructure, enabling researchers to map the participation of different population groups in a local social space of various scenes (Otte 2004, chap. 11).

Related to these communal citizens’ surveys are audience and visitor surveys borne by cultural institutions like museums and theatres. Here we expect even greater variation in data quality. A careful methodological assessment should be made before using data for secondary analyses. This survey approach is insightful because the composition of the audiences consuming concrete aesthetic products and services can be studied on the basis of actual (not reported) behavior. Such data enrich aggregate visitor statistics that are reported by cultural institutions and professional associations. Informative spatial-temporal comparisons are enabled by combining various audience samples (Dollase et al. 1986; Klein 1990; Rössel et al. 2005). Unfortunately, documentation of such studies is even scarcer and access to datasets more problematic.

Finally, we would like to mention three more sources of individual-level survey data which could be usefully employed for scientific analyses. First, official statistics, such as the Sample Survey of Income and Expenditure (EVS, *Einkommens- und Verbrauchsstichprobe*), the Household Budget Survey (LWR, *Laufende Wirtschaftsrechnungen*), and the Time Use Survey of the Federal Statistical Office, do not sufficiently differentiate cultural consumption activities and expenditures internally. Instead, they tend to merge “culture” and “leisure” categories. These categorizations could be improved. Second, the media research departments of the public radio stations, ARD and ZDF, carry out nation-wide studies (e.g., “ARD/ZDF-Onlinestudie,” “ARD-E-Musik-Studie,” surveys employing the “MedienNutzerTypologie”) and even more studies confined to single transmission areas on various

---

14 Susanne Plagemann, responsible for documentation issues at the Difu, gave us rich information about komm.DEMOS. It is accessible on a “fee for service” basis via the IRB Stuttgart ([www.irb.fraunhofer.de/datenbanken.jsp](http://www.irb.fraunhofer.de/datenbanken.jsp)). The study descriptions contain information on the primary researcher who might be asked for the release of survey data for secondary analyses. Where ever local statistical agencies collected the data, chances are great that the data are still existent. Only in exceptional cases were they given to the GESIS Data Archive. Rudolf Schulmeyer, chairman of the VDSt, promised to put our request about the trans-communal data infrastructure on the agenda of the next executive board meeting.



aspects of media consumption and musical preferences. Only a few of these data have been made accessible for scientific secondary analyses. In particular, the older data could be placed at the GESIS Data Archive's disposal, like those of the Leser- and Media-Analyse recently have been (Hagenah et al. 2006).<sup>15</sup> Third, cultural preferences and activities are frequently part of market research surveys. Some have been given to the GESIS Data Archive, such as, "Outfit 1-4" or "Typologie der Wünsche," but many more could be made available.

#### 4. Recommendations

Taking into account the research needs in the sociology and economics of culture, the status quo of data infrastructure, and current debates in official statistics and cultural policy, we conclude with the following recommendations:

- (1) A double-task of prime importance that has to be accomplished by scholars in academia is the theory-driven development of two comprehensive, large-scale "baseline" surveys. The first one has to follow the "social-status-of-artist" and "Künstlerreport" tradition, but should also contain detailed life-course information enabling analyses of artists' careers. The second one has to be a representative population survey on cultural consumption and reception comprising current preferences and behavior. Additionally, retrospective biographical and social network information should be included. These surveys call for public funding (e.g., by the DFG). They can serve as baselines for the construction of more elaborate panel studies on culture, as well as replications in an international or European comparative context.
- (2) We support the inquiry into "Culture in Germany" (*Kultur in Deutschland*) in its recommendation of the construction of a nationally unified and standardized cultural statistic, mainly based on aggregate data, borne by the Federal Statistical Office and compatible with efforts at the EU level. It should allow researchers to distinguish at least between the core of the cultural sector and a wider notion of the cultural industries (KEA European Affairs et al. 2006), between the public, private, and non-profit sectors, and different cultural branches. For adequate scientific research, differentiated data on low aggregation levels are needed.

---

15 We contacted Dr. Ekkehardt Oehmichen, director of media research at the Hessischer Rundfunk, who promised to address this topic at the next meeting of ARD media researchers.

- (3) Organizational-level data, especially those collected for the theater and museum statistics, should be made available in a computer-readable format in order to facilitate statistical analyses.
- (4) The large pool of communal citizen surveys on cultural topics and of organization-based audience surveys should be documented and made accessible in a central archive. Three options seem to be practicable. First, the German Institute of Urban Affairs' (Difu, *Deutsches Institut für Urbanistik*) database "komm.DEMOS" should be financially supported in order to enable scientific access free of charge. We recommend this step even if the database is not expanded to cover survey datasets. Further funding would enable archival storage and administration of such data at the Difu. Second, the GESIS Data Archive for the Social Sciences, with its approved data infrastructure, could be an alternative archival location. Third, a Research Data Center for data of communal statistical agencies could be established at the Federal Statistical Office. In all cases alike, studies should be carefully selected and documented according to scientific requirements of data quality.
- (5) Access to data on culture collected by statutory bodies (media research of public radio stations), by the Centre for Cultural Research (*Zentrum für Kulturforschung*) (e.g., "Kulturbarometer," "Künstlerreport"), and by market research institutes, should be improved. These data are promising for building up time-series and analyzing trends in cultural preferences and behavior. The GESIS infrastructure would be suited best as an archive for these data. Conversion of data from the 1970s into electronically readable files would also be worthwhile funding if data quality is satisfying and studies are important for historical-comparative work.

Neither in Germany nor abroad is the field of culture well-institutionalized in its current research infrastructure. The field is of growing importance, though, not only in the social and economic sciences, but also in society and the economy in general. Improving data access and supporting large-scale surveys would assist scholars in Germany greatly in their effort to reach a leading international research position in this thriving field.

## References:

- Baumann, Sh. (2007): *Hollywood Highbrow. From Entertainment to Art*. Princeton.
- Becker, H.S. (1982): *Art Worlds*. Berkeley.
- Beckert, J. and Rössel, J. (2004): Kunst und Preise. Reputation als Mechanismus der Reduktion von Ungewissheit auf dem Kunstmarkt. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 56, 32-50.
- Beyers, T. (2005): Cultural Education and the Canon. A Comparative Analysis of the Content of Secondary School Exams for Music and Art in England, France, Germany and the Netherlands, 1990-2004. *Poetics* 33, 388-416.
- Blau, J.R. (1988): Study of the Arts: A Reappraisal. *Annual Review of Sociology* 14, 269-292.
- Blaug, M. (2001): Where are we now on Cultural Economics? *Journal of Economic Surveys* 15, 123-143.
- Bourdieu, P. (1999): *Die Regeln der Kunst. Genese und Struktur des literarischen Feldes*. Frankfurt a.M.
- Bourdieu, P. and Darbel, A. with assistance of Schnapper, D. (2006): *Die Liebe zur Kunst. Europäische Kunstmuseen und ihre Besucher*. Konstanz.
- Bretschneider, M. with assistance of Schumacher, J. (1996): DEMOS – Eine Datenbank zum Nachweis kommunaler Umfragen auf dem Weg zum Analyseinstrument. *ZA-Information* 38, 59-75.
- Caves, R.E. (2000): *Creative Industries. Contracts between Art and Commerce*. Cambridge.
- Deutscher Bühnenverein (2008a): *Theaterstatistik 2006/2007*. Deutschland. Österreich. Schweiz. Köln: Deutscher Bühnenverein, Bundesverband der Theater und Orchester.
- Deutscher Bühnenverein (2008b): *Wer spielte was? Werkstatistik 2006/2007*. Deutschland. Österreich. Schweiz. Bensheim.
- Deutscher Bundestag (2007): *Schlussbericht der Enquete-Kommission „Kultur in Deutschland“*. Berlin: Deutscher Bundestag, Drucksache 16/7000.
- Deutscher Städtetag (1994): *Methodik von Befragungen im Kulturbereich. Eine Arbeitshilfe zur Planung, Durchführung, Auswertung und Präsentation von Besucherbefragungen in Museen und Theatern sowie von Befragungen über Kultur in allgemeinen Umfragen / Bürgerbefragungen*. Köln: Reihe H-DST-Beiträge zur Statistik und Stadtforschung, Heft 40.
- DiMaggio, P. and Mukhtar, T. (2004): Arts Participation as Cultural Capital in the United States, 1982-2002: Signs of Decline? *Poetics* 32, 169-194.
- Dollase, R./Rüsenberg, M. and Stollenwerk, H.J. (1986): *Demoskopie im Konzertsaal*. Mainz.
- Dowd, T. (2004): Concentration and Diversity Revisited. Production Logics and the U. S. Mainstream Recording Market, 1940-1990. *Social Forces* 82, 1411-1455.
- Eurostat (2007): *Cultural Statistics*. Luxembourg.
- Fohrbeck, K. and Wiesand, A.J. (1972): *Der Autorenreport*. Reinbek.
- Fohrbeck, K. and Wiesand, A.J. (1975): *Der Künstler-Report. Musikschaffende, Darsteller/Realisatoren, Bildende Künstler/Designer*. Munich.
- Frey, B.S. (2000): *Arts and Economics: Analysis and Cultural Policy*. Berlin.

- Gans, H.J. (1974): *Popular Culture and High Culture. An Analysis and Evaluation of Taste*. New York.
- Ganzeboom, H.B.G. (1982): Explaining Differential Participation in High-Cultural Activities. A Confrontation of Information-Processing and Status-Seeking Theories. In: Raub, W. (Ed.): *Theoretical Models and Empirical Analyses. Contributions to the Explanation of Individual Actions and Collective Phenomena*. Utrecht.
- Gebesmair, A. (2008): *Die Fabrikation globaler Vielfalt. Struktur und Logik der transnationalen Popmusikindustrie*. Bielefeld.
- Gerhards, J. (1997): Soziologie der Kunst: Einführende Bemerkungen. In: Gerhards, J. (Ed.): *Soziologie der Kunst. Produzenten, Vermittler, Rezipienten*. Opladen.
- Gerhards, J. (2008): Die kulturell dominierende Klasse in Europa. Eine vergleichende Analyse der 27 Mitgliedsländer der Europäischen Union im Anschluss an die Theorie von Pierre Bourdieu. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 60, 723-747.
- Ginsburgh, V.A. and Throsby, D. (Eds.) (2006): *Handbook of the Economics of Art and Culture*. North-Holland.
- Giuffre, K. (1999): Sandpiles of Opportunity. Success in the Art World. *Social Forces* 77, 815-832.
- Greenfeld, L. (1988): Professional Ideologies and Patterns of „Gatekeeping“. Evaluation and Judgement in Two Art Worlds. *Social Forces* 66, 903-925.
- Haak, C. (2008): *Wirtschaftliche und soziale Risiken auf den Arbeitsmärkten von Künstlern*. Wiesbaden.
- Hagenah, J./Meulemann, H. and Akinci, H. (2006): *German Media-Analyse (MA). A Large-Scale Commercial Data Source Available for Secondary Analyses on Media Use and Social Change*. *Schmollers Jahrbuch* 126, 129-137.
- Hartmann, P.H. (1999): *Lebensstilforschung. Darstellung, Kritik und Weiterentwicklung*. Opladen.
- Hügel, H.-O. (Ed.) (2003): *Handbuch Populäre Kultur. Begriffe, Theorien und Diskussionen*. Stuttgart/Weimar.
- Institut für Museumsforschung (2007): *Statistische Gesamterhebung an den Museen der Bundesrepublik Deutschland für das Jahr 2006*. Berlin: Staatliche Museen zu Berlin – Preußischer Kulturbesitz, Institut für Museumsforschung.
- Janssen, S. (2001): The Empirical Study of Careers in the Literature and the Arts. In: Schram, D. and Steen, G. (Eds.): *The Psychology and Sociology of Literature*. Amsterdam.
- Jonsson, J.O./Grusky, D.B./Di Carlo, M./Pollak, R. and Brinton, M.C. (2007): *Micro-Class Mobility. Social Reproduction in Four Countries*. Mannheim: MZES Working Paper No. 100.
- Karttunen, S. (1998): How to Identify Artists? Defining the Population for “Status-of-the-Artist” Studies. *Poetics* 26, 1-19.
- Katz-Gerro, T./Raz, Sh. and Yaish, M. (2007): Class, Status, and Intergenerational Transmission of Musical Tastes in Israel. *Poetics* 35, 152-167.
- KEA European Affairs, Turku School of Economics and MKW Wirtschaftsforschung (2006): *The Economy of Culture in Europe. Study prepared for the European Commission, Directorate-General for Education and Culture*. Brussels.
- King, T. (2007): Does Film Criticism Affect Box Office Earnings? *Journal of Cultural Economics* 31, 171-186.

- Kirchberg, V. (2005): Gesellschaftliche Funktionen von Museen. Makro-, meso- und mikrosoziologische Perspektiven. Wiesbaden.
- Klein, H.J. (1990): Der gläserne Besucher. Publikumsstrukturen einer Museumslandschaft. Berlin.
- Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik (KVI) (Ed.) (2001): Wege zu einer besseren informationellen Infrastruktur. Baden-Baden.
- Kressin, J. (2008): Zwischen Autonomie und Abhängigkeit. Empirische Analysen zur Marktsituation selbständiger Künstler in Deutschland. Leipzig: Universität Leipzig, Institut für Kulturwissenschaften (Magisterarbeit).
- Lamont, M. and Molnár, V. (2002): The Study of Boundaries in the Social Sciences. *Annual Review of Sociology* 28, 167-195.
- Mayer, K.U. (2008): Retrospective Longitudinal Research: The German Life History Study. In: Menard, S. (Ed): *Handbook of Longitudinal Research: Design, Measurement, and Analysis*. Burlington.
- Mark, D. (1998): Wem gehört der Konzertsaal? Das Wiener Orchesterrepertoire im internationalen Vergleich. Wien/Mülheim.
- Menger, P.-M. (1999): Artistic Labor Markets and Careers. *Annual Review of Sociology* 25, 541-574.
- Otte, G. (2004): Sozialstrukturanalysen mit Lebensstilen. Eine Studie zur theoretischen und methodischen Neuorientierung der Lebensstilforschung. Wiesbaden.
- Otte, G. (2008): Lebensstil und Musikgeschmack. In: Gensch, G./Stöckler, E.M. and Tschmuck, P. (Ed.): *Musikrezeption, Musikdistribution und Musikproduktion. Der Wandel des Wertschöpfungsnetzwerks in der Musikindustrie*. Wiesbaden.
- Peterson, R.A. (Ed.) (1976): *The Production of Culture*. Beverly Hills.
- Peterson, R.A. (2005): Problems of Comparative Research. The Example of Omnivorousness. *Poetics* 33, 257-282.
- Rössel, J./Hackenbroch, R. and Göllnitz, A. (2005): Soziale Differenzierung und Strukturwandel des Hochkulturpublikums. In: Institut für Kulturpolitik der Kulturpolitischen Gesellschaft (Ed.) (2005): *Jahrbuch für Kulturpolitik 2005*. Essen.
- Schneider, M.A. (1993): *Culture and Enchantment*. Chicago.
- Schulze, G. (1992): *Die Erlebnisgesellschaft. Kultursoziologie der Gegenwart*. Frankfurt.
- Shrum, W.M. (1997): *Fringe and Fortune. The Role of Critics in High and Popular Art*. Princeton.
- Simonton, D.K. (1997): Creative Productivity: A Predictive and Explanatory Model of Career Trajectories and Landmarks. *Psychological Review* 104, 66-89.
- Söndermann, M. (2005): Beschäftigung im Kultursektor in Deutschland 2003/2004. Ergebnisse der Kulturstatistik. In: Institut für Kulturpolitik der Kulturpolitischen Gesellschaft (Ed.) (2005): *Jahrbuch für Kulturpolitik 2005*. Essen.
- Statistische Ämter des Bundes und der Länder (Eds.) (2008): *Kulturstatistiken. Kulturindikatoren auf einen Blick. Ein Ländervergleich*. Wiesbaden: Hessisches Statistisches Landesamt.
- Statistisches Bundesamt (Ed.) (1994): *Im Blickpunkt: Kultur in Deutschland. Zahlen und Fakten*. Stuttgart.

- Statistisches Bundesamt (2004): Methodenkritische Analyse von Basisstatistiken zum Kulturbereich und Fragen zu einem Anforderungsprofil an eine bundeseinheitliche Kulturstatistik für die Bundesrepublik Deutschland. Gutachten für die Enquete-Kommission „Kultur in Deutschland“ des Deutschen Bundestages. Langfassung. Wiesbaden.
- Throsby, D. (1994): The Production and Consumption of the Arts: A View of Cultural Economics. *Journal of Economic Literature* 32, 1-29.
- Thurn, H.P. (1983): Die Sozialität der Solitären. Gruppen und Netzwerke in der Bildenden Kunst. In: Neidhardt, F. (Ed.): *Gruppensoziologie. Perspektiven und Materialien*. Opladen.
- Weckerle, Ch./Söndermann, M. and Hochschule für Gestaltung und Kunst Zürich (2003): Das Umsatz- und Beschäftigungspotential des kulturellen Sektors. Erster Kulturwirtschaftsbericht Schweiz. Zurich.
- Weishaupt, H. and Fickermann, D. (2001): Informelle Infrastruktur im Bereich Bildung und Kultur. Expertise für die Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik. CD-Rom Attachment in: Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik (KVI) (Ed.) (2001): *Wege zu einer besseren informationellen Infrastruktur*. Baden-Baden.
- Zolberg, V. (1990): *Constructing a Sociology of the Arts*. Cambridge.

## 6.4 Mass Media Research

*Heiner Meulemann and Jörg Hagenah*

Contact:

Heiner Meulemann  
Jörg Hagenah  
Forschungsinstitut für Soziologie  
Universität zu Köln  
Greinstraße 2  
50939 Cologne  
Germany  
e-mail: meulemann[at]wiso.uni-koeln.de  
hagenah[at]wiso.uni-koeln.de



## Abstract

Mass media are defined as media that have their own, self-defined program and solicit their own audience. Accordingly, mass media research deals with the production of programs and the consumption of programming by an audience. This report discusses the justification for both perspectives on the topic, introduces the relevant data sources, and makes recommendations for the research infrastructure. In terms of media production, the discussion recommends the establishment of a central media content archive where content analytic time series of public agencies as well as the work of individual researchers can be collected. Furthermore, the development of a unified system of content analysis and the promotion of cross-national comparisons are recommended. In terms of media consumption, the provision of privately funded data to the scientific community, the promotion of cross-national comparisons, and the linkage of programs and audience data are recommended.

Keywords: mass media, data archive, content analysis, survey research

## 1. Introduction

Media can be defined as a set of technologies designed to store and distribute meaning. Among media in general, *mass* media can be singled out by both the types of meaning it produces and by the audience receiving these messages. In terms of meaning, mass media content is produced by specialized agencies according to a predetermined schedule of “(daily) actuality” within a particular national or linguistic community (Reitze and Ridder 2006). Mass media have a program: they pre-package content and distribute it according to a thematic structure and a time schedule – they are media for masses of meanings. Regarding the audience, the use of mass media is delineated by the technical requirements and possibilities of mass media themselves, together with the given language, so that in principle they are available for each member of a nation or language community, rather than for socially circumscribed groups only – they are media for masses of people.

Mass media, therefore, can be distinguished from individual media, such as the book, the letter, the telephone, and the internet. The meanings of individual media are produced by persons individually; they are derived by individuals according to personal needs and have a small, socially restricted audience, such as friends, the family, and professional or intellectual peers, audiences that often can be named, for example, the “intellectuals” or the “*Bildungsbürgertum*.” Mass media are anchored in a national society; individual media are anchored in – as internet jargon has it – “communities” that rest on personal, although not face-to-face, relationships.

As mass media address nations, their development is one strand in the modernization of nations (Hallin and Mancini 2004: 261). Up to now, modern societies have only had three forms of mass media: (daily) newspapers (including periodicals), dating back to the seventeenth century; radio, originating in the 1920s; and television, taking the lead after the 1950s. These three will be dealt with in the following expert report. The internet, however, will not be dealt with, since it is an individual rather than a mass medium. It is a technical platform, which is primarily used for personal communication and for personal services, but which can also be used in order to distribute the three above-named mass media (Meulemann 2009). Therefore, it has been identified as a “converged medium” (PriceWaterhouse-Coopers 2007: 53).

If mass media have a program and solicit their own audience, mass media research comprises both a communicator-oriented perspective on the production of programs and a recipient-oriented perspective on consumption by audiences. In the following, we describe the research topics of both perspectives in section 2, their data sources in section 3, and give recommendations for the research infrastructure of both in section 4. For the sake of simplicity, we speak of media only although we refer to mass media throughout.

## **2. Research topics and research questions**

### *2.1 Production of programs*

The appropriate method to analyze the production of media programming is content analysis. Its topics can be broadly grouped, as one view from media studies journals explains (Bonfadelli 2002: 33), into two classes: *analyses of information structure* and *analyses of social problem areas*. The former purport to examine whether the media fulfill their social function as a “fourth public authority” and satisfy the information needs of the audience. The latter intends to examine whether the media discriminate against social groups or represent them adequately.

#### Analyses of information structure

The core questions treated in this research area concern whether and how the introduction of the dual broadcasting system in the middle of the 1980s has marginalized public stations or assimilated them into new private ones. The *marginalization hypothesis* contends that public stations have lost audience share to private stations from the lower social strata (Krüger 1992). As a

result of this loss, public broadcasting might suffer from a legitimacy crisis concerning their “public mandate of a basic provision” (Bomas 2005) or their “cultural mandate” (Rossen-Stadtfeld 2005) as well as their mandatory dues system (Kleinstauber et al. 1991). The *convergence hypothesis* contends that the public stations tend to neglect their “public mandate” in devoting an ever larger part of their program to mass-entertainment (Brosius and Zubayr 1996, 186), and that private stations have improved their informational broadcasting in order to catch up in this area to the public stations (Saxer 1980; Schatz et al. 1989).

Numerous content analyses, mostly concerned with television and seldom with radio (Marchal 2004: 704ff), have examined these two hypotheses (Maier 2002: 83; Brosius and Zubayr 1996). However, the results have remained contradictory. On the one hand, Krüger (Krüger and Zapf-Schramm 2008; Krüger 2005b; Krüger 1992; 2001; Krüger and Zapf-Schramm 2002) has compared genres, broadcasts, and contents of public and private stations annually since 1985 and regularly detected differences between both groups of channels on all three levels (Weiß 2007; Trebbe 2004; Trebbe and Weiß 1994: 175; Meier 2003). On the other hand, assimilative tendencies between public and private stations in the program structure (Faul 1988; 1989; Schatz et al. 1989; Donsbach and Dupré 1995; Hallermann et al. 1998; Sutor 1999; Rossmann et al. 2003), in the presentation of newscasts (Kaase 1989; Pfetsch 1991; Greger 1998; Goertz 1996), and in sports reporting (Scholz 1993) have also been shown to exist.

Apart from this core question, content analyses about daily newspapers have only been used in a few stand-alone studies (Meier 2002: 192; Maurer and Reinemann 2006: 83). These studies are hardly comparable, because they refer to different titles and use different content analytical categories. However, a few tendencies relating to the content areas of politics, economics, and sports can be summarized. Thus, in all newspapers examined – with the exception of the tabloid “Bild” (Schulz 1970) – politics and economics have dominated since the 1950s, and sports have captured a considerable portion of the news content in each newspaper (Held and Simeon 1994; Hüther et al. 1973; Schulz 1970; Schwantag 1974; Hagemann 1958).

#### Analyses of social problem areas

The question posed in this category of research concerns whether social groups are adequately reflected in media content. In the simplest case, this has been done by comparing the shares of groups represented in the media with their share in the population (Bonfadelli 2002, 33ff), particularly the share of foreigners (Bonfadelli 2007; Bonfadelli and Moser 2007; Ruhrmann 2002), of poor and old people (Burgert and Koch 2008; Davis and Kubey 1982; Bosch 1988), and of men and women (Gnändiger 2007; Petersen 2006;

Hesse 2001; Jud-Krepper 1997; Fröhlich and Holtz-Bacha 1995; Werner and Rindsdorf 1998; Weiderer 1992; Schmerl 1985; Ulze 1977; Küchenhoff 1975) have been compared.

More complex studies of media representation examine specific social problems. Thus, the presentation of crime and violence has been investigated with respect to their possibly detrimental effects on social integration (Kunczik 2008; Petzold 2008; Gerbner et al. 1979; 1980; Stein-Hilbers 1977; Groebel and Gleich 1993). In a similar vein, reports on racism (Handel 1998; Ruhrmann and Kollmer 1987), on conflicts and wars (Fröhlich et al. 2007; Kolmer 2004; Hallin 1997; Olien et al. 1989), on drug abuse (Fleming et al. 2004; Rose 1995), and on pornography (Scheufele 2005; Brosius and Rössler 1999; Amann and Wipplinger 1997) have been subjected to content analysis.

## 2.2 Media consumption

In order to investigate the media audience within a particular country, nationally representative population samples have to be surveyed. This results in a *micro perspective* which examines persons which use and do not use mass media and why (Lindner-Braun 2007; Meyen 2004; Schweiger 2007). If replicated, they also provide information relevant to a *macro perspective* on media systems and their development.

### Micro perspective

As research on media consumption is initially driven by the needs of advertising research (*Werbeträgerforschung*), it begins with (1) the *socio-demography* of media use that allows producers to find their audiences and to calculate their advertising price. This very purpose, moreover, requires frequent replications of the surveys with the same question wording. The socio-demographic variables thus surveyed encompass the basic opportunities and restrictions of media use, such as the resources of education, occupation, and income and the obligations associated with being employed and having a family. Thus, the relative impact of these factors on media use can be studied. However, advertising research has already gone well beyond that in collecting information on (2) *time budgets* of work, leisure, and media use, so that media use can be related to its most important resource – leisure time. Thus, it is possible to examine how all these resources determine media use; moreover, this impact can be followed up over time.

Nevertheless, media use – like all forms of consumption – depends not only on resources but on (3) *preferences* as well. These determine how time is allocated to work or leisure, to indoor or outdoor leisure activities, and to competing media sources. But preferences for information or entertainment

or for political and fictional contents in the media, for active or passive leisure pastimes in general, or for leisure or work are never investigated in advertising research and only rarely in academic research.

Finally, media use depends on (4) *attachments* to and (5) *images* of specific media. Questions such as how much one would miss a specific medium, or how trustworthy a specific medium is assessed to be, have been repeatedly asked in both advertising and academic research (Reitze and Ridder 2006: 26-32, 80-95).

#### Macro perspective

Taken as a whole, replicated surveys delineate changes in media use and in the underlying social structure. There are two ways to examine the relation between these two levels of change. First, once social structural developments have been controlled for, the (1) *total and net change of media use* can be compared. For example, Fürtjes (2008) examined whether the changing composition of a German soccer fan magazine between 1954 and 2005 reflected concurrent changes among the media or in the social structure, and demonstrated that only the latter was responsible for changes in the readership. The phenotypic change disappeared once changes in population composition had been controlled for; there were no genuine media developments in this domain beyond social structural changes. Moreover, changes in media use can be at least partly (2) *explained by cohort succession*. Thus, the cohort that first experienced television might become “the television cohort” and stick more than other cohorts to television viewing during their lifetimes – which was indeed examined and did not turn out to be true (Peiser 1996). Similarly, the cohorts that first experienced private television might stick more than others to private programs – which indeed did turn out to be true, but which only partly explained the audience movements from public to private broadcasters (Meulemann et al. 2009).

In addition, since Germany, like all other European countries, switched during the 80s from a monopoly of public broadcasting to a dual broadcasting system, (3) the *effects of organizational change on media use* can be examined within the total audience as well as specific segments.

Finally, as advertising research often contains information on the use of a whole range of media used by a person, changes in the (4) *media repertory* as well as in the encompassing *consumer repertory* can be assessed. Is the increasing number of television broadcasters and the decreasing number of newspapers reflected in corresponding changes in personal repertoires? Similar questions can be treated on the aggregate level: which media gain at the cost of which others? Which media compete within a specific market of, say, periodicals, which substitute for each other?

If content analyses and population surveys are combined, the perspective can be broadened from communication to social research in general – as illustrated in the case of two topics: leisure and politics. In studying leisure, the combination of media content analyses and time budget surveys of media use allows for a more convincing investigation into why people prefer the media to other leisure pastimes, and in particular why the preference for the media, as measured by their share of leisure time, has increased in Germany after the introduction of the dual system, although the supply for other leisure time activities, such as theaters and museums, clubs and associations, has risen simultaneously as well (Gilles et. al. 2008). As for politics, the combination of media content analyses with surveys on the perception of politics and politicians and on voting decision allows for an investigation of media effects on the political process (Petersen and Jandura 2006; Semetko 2009). The “political communication” (Schulz 2003) between citizens, media and politicians could be followed up, not only on the aggregate level, but also on the individual level.

### 3. Sources for media research

#### 3.1 *Production of programs*

##### Archiving institutions

In principle, all mass media can be archived continuously so that they can be content analyzed even decades after they have been issued or broadcasted. In practical terms, however, print and electronic media are accessed in different ways.

In the case of *print media*, supra-regional and above all also regional daily papers are archived in municipal libraries or city archives that want to keep track of their own history. Beyond this, a copy of each paper must be held by the German National Library (*Deutsche Bibliothek*) in Frankfurt. For some newspapers, even digital versions are provided by the internet software of the commercial info service LexisNexis.<sup>1</sup>

As for *electronic media*, broadcasting contents are less systematically archived and therefore less easily accessible. The public stations archive their program galleys in the German Broadcasting Archive (*Deutsches Rundfunkarchiv*) in Frankfurt and their self-produced broadcasts in archives of the ARD and ZDF stations. In analyzing these archives two content analytic strategies have been followed. Either current programs are video- or DVD-

---

1 [www.lexisnexis.de](http://www.lexisnexis.de)

taped or the program structure published in program magazines is analyzed (Merten 1996: 156). The first strategy permits broad investigations of current programs, but allows no longitudinal designs. The second strategy permits longitudinal analyses, yet has some shortcomings, since published programs are changed in the short run to give space to unforeseen events and do not allow deeper analyses with fine-tuned categories.

Using these archives, time series which refer to the program structure of television and radio in general and to news broadcasts of the television specifically have been constructed.

Time series: The program structure of television and radio in general

With regard to the structure of television programs, national<sup>2</sup> time series are regularly constructed by four research groups. (1) the Institute for Empirical Media Research (IFEM, *Institut für empirische Medienforschung*) has been tracking the genre profiles of public stations *ARD* and *ZDF* and private stations *RTL*, *SAT.1* and *Prosieben* since 1985 annually on the basis of four broadcasting weeks (Krüger 2005a: 302; Krüger and Zapf-Schramm 2008). These analyses aim at examining whether there is a convergence of public and private programming or not. (2) GöfaK Media Research (*GöfaK Medienforschung*) analyzes most of television programs since 1998 on behalf of the Association of State Media Authorities for Broadcasting in Germany (ALM, *Arbeitsgemeinschaft der Landesmedienanstalten*). One broadcasting week of the public stations *ARD* and *ZDF* and of the private stations *RTL*, *RTL II*, *Vox*, *Sat.1*, *Prosieben* and *kabel 1* are videotaped each spring and each autumn (Weiß 2007; Trebbe 2004; Weiß and Trebbe 2000; Weiß 1999). These analyses aim at giving private stations some feedback about their success within the dual broadcasting system. (3) The most elaborate analysis is commissioned by the Consortium for TV Research (AGF, *Arbeitsgemeinschaft Fernsehforschung*). Since 1963, the telemetric data – at that time still only of *ARD* and *ZDF* – have been linked with program data such that social profiles of the program use can be established. Since 1985, these measurements, which as of today cost 17 million Euros per year, are performed by the Society for Consumer Research (GfK, *Gesellschaft für Konsumforschung*); Hagenah and Meulemann 2007: 157f). The data are collected to calculate advertisement prices in specific program areas. (4) Using a different source, namely the television listings magazine *Hörzu* from 1980 to 1993, Merten (1994; 1996), commissioned by the German Association of Private Broadcasters and Telemedia (VPRT, *Verband Privater Rundfunk und Telemedien*), analyzed the content of eighteen public and

---

2 Furthermore, there are stand-alone studies of regional television programs, e.g., in Thüringen 1999, 2002 and 2006 (tlm.de 2000; Moses and Heyen 2003; Giewald and Heyen 2007).

private television stations. The analyses aimed at showing that private full programs<sup>3</sup> provided information in a manner comparable to the public stations.

As for the structure of *radio* programs, time series that compare several *Länder*<sup>4</sup> are regularly constructed by three research groups. (1) The year-books of the *ARD* report the percentages of word and music broadcasts from all *ARD* stations since 1969 (Gleich 1995: 555). Using this source, each of the *ARD* stations has analyzed the content of its programming extensively; unfortunately, they have changed their analytical categories between time points. (2) Wichert (2008; 1997) has examined the content profiles of private radio programs in comparison with public programming in Berlin-Brandenburg for the State Media Authority of Berlin-Brandenburg (*Landesmedienanstalt Berlin-Brandenburg*) since 1994. (3) Heyen (2001) has examined the program structure of *Antenne Thüringen*, *Landeswelle Thüringen*, *Jump*, and *MDR 1 Radio Thüringen* on behalf the State Media Authority of Thuringia (*Landesmedienanstalt Thüringen*) since 1996.

#### Time series: News broadcasts of television

As for the structure of the *news broadcast*, time series have been constructed by four research groups. (1) The *InfoMonitor* of the IFEM institute has examined all the main newscasts of *ARD*, *ZDF*, *RTL*, and *Sat.1* according to the structure of their topics on behalf of these public stations since 2005 (Krüger 2008; 2005b). (2) Maier, Ruhrmann and Klietsch (2006) have analyzed the topic structure for the main newscasts of *ARD*, *ZDF*, *RTL*, *RTL II*, *ProSieben*, *Vox*, and *kabel 1* on behalf of the State Media Authority of North Rhine-Westphalia (*Landesmedienanstalt Nordrhein-Westfalen*) at five time points between 1992 and 2004. (3) The Institute for Media Research (IMGÖ, *Institut für Medienforschung*) examines the regional news reporting in the *Länder* of former West Germany broadcasted by *RTL* and *Sat.1* on behalf of the ALM annually since 2005 (Volpers et al. 2008).

(4) The private institute *Media Tenor* has constructed the most encompassing content analyses since 1993. It has no commissioning agency, but sells its analyses to enterprises who are interested in gaining knowledge about their representation in the media. It scans approx. 700 media every day worldwide – among them the most important German television and radio

3 Full programmes offer a programm-mix, they are not only specialised in themes like music, entertainment, news or sports.

4 There are also stand-alone studies of regional radio broadcasting programs that have been commissioned by the media authority of the *Länder* (*Landesmedienanstalten*), for example in Schleswig-Holstein (Hasebrink 2006), Niedersachsen (Volpers 2009), Mecklenburg-Vorpommern (Rager et al 2005), Nordrhein-Westfalen (Volpers and Schnier 1996), Hessen (Brosius and Weiler 2003), Rheinland Pfalz (Rager and Siebers 2006), Saarland (Bauer 2003), Baden-Württemberg (Schönbach et al. 1993), and Bayern (Stuiber 1990).



stations, all supra-regional, and the most important regional newspapers, the news magazines, and the most important news sites and news blogs available on the Internet (mediatenor.de 2009).

### 3.2 Audience media consumption

#### Surveys commissioned by broadcasting agencies

Three large-scale sources provide answers to the questions about audience described in section 2.2. (1) Every half decade from 1964 to 2005 the *ARD-Werbung Sales & Services* has commissioned the Long-Term Study on Mass Communication (*Langzeitstudie Massenkommunikation*), which surveys the use of the daily media (newspapers, radio, and television) and the internet, the time budgets of media use, audience attachment, and the image of the media (Reitze and Ridder 2006). This source contains answers to nearly all the questions relevant to a micro perspective. In terms of macro perspective, its datasets cumulatively describe the often dramatic changes in media use (Schweiger 2007: 42-48), such that it can be explained by cohort succession (Peiser 1996; Engel and Best 2001; Reitze and Ridder 2006: 134-165).

(2) Since 1954 the Consortium of Media Analyses (AG-MA, *Arbeitsgemeinschaft Media-Analyse*) has commissioned the Reader Analysis (LA, *Leser-Analyse*), which since 1972 has been called the Media Analysis (MA, *Media-Analyse*). They survey the use of many print media titles and electronic media stations minutely through a series of question modules in order to establish the “advertising currency” of each. Since 1987, they have also administered a time budget question module (Hagenah and Meulemann 2009a; b). In order to do this for specific titles, samples have to be very large (more than 60,000). To detect general developments in media use, however, titles and programs have to be regrouped into genres. For example, periodicals have to be grouped into title categories such as politics, sports, etc.; and radio programs must be grouped into tune in and go along programs (*Einschalt- und Begleitprogramme*), predominantly devoted to information or to entertainment. The Center for Teaching and Research in Media Science (*Medienwissenschaftliches Lehr- und Forschungszentrum*) has accomplished this sometimes extensive work, together with the technical preparation of the datasets. In addition, it has constructed time series of the use of all types of media and of social structures from the MA datasets, which are available on their website.<sup>5</sup> Thus, the LA and MA present many options to researchers for describing changes in media use in the context of changing social structure from a macro perspective. Since quite a few of the MA even contain a

---

5 <http://www.mlfz.uni-koeln.de/index.php?id=106>.

question on party preference, they also can also be used for political analyses.

(3) The MA discontinued its survey of the use of specific television stations in 1997 and only continues to report the total television time. Since then, only telemetric information is available for specific stations. Telemetric research started in Germany already in 1963, commissioned by the public stations. At present, the AGF has commissioned the Society for Consumer Research (AGF/GfK, *Gesellschaft für Konsumforschung*) to measure for every second of television use by each member of a panel of households telemetrically in order that results are ready the day after broadcasting (Lindner-Braun 2007: 127-139). These data – as well as their predecessors – are not available to the public. However, on request the AGF delivers the reaches and the market participation of several age brackets of each broadcasting for each station immediately in the following week.

Apart from these studies of media use in general, there are three further kinds of studies partly or completely commissioned by broadcasting agencies that either put media use in the broader perspective of consumption and survey the total population (4-6) or that put media use in the deeper perspective of the evaluation of programs and survey restricted populations, either regionally (7) or according to age (8).

(4) The Consumer and Media Use Analysis (VUMA, *Verbrauchs- und Medienanalyse*)<sup>6</sup> research media use in the context of general consumer behavior in order to detect complementarities and substitutions within each and between both. The samples are nationally representative and somewhat smaller than those of the MA (about 24,000). VUMA started in 1995 and has been replicated annually since 2000. The same aim is served by (5) the Allensbach Market and Advertising Media Analysis (AWA, *Allensbacher Markt- und Werbeträger Analyse*)<sup>7</sup> and by (6) the survey, Typology of Needs (TDW, *Typologie der Wünsche*)<sup>8</sup>. AWA started in 1959 and has been replicated annually with samples of about 20,000. TDW started in 1974 and has been replicated annually since 1986 with samples of about 20,000.

(7) For quite a few of the German *Länder*, *state specific MA studies* on the use of radio and television have been administered. They have their own names, such as the Bavarian Broadcasting Analysis (*Funkanalyse Bayern*)<sup>9</sup> that started in 1989 and is replicated annually and comprises samples of about 23,000 for radio use and a further 16,000 for television use. These studies have two primary goals. First, they provide data on small and local stations. Second, and more interesting for academic purposes, they survey the

---

6 [www.vuma.de](http://www.vuma.de).

7 [www.aw-online.de](http://www.aw-online.de).

8 [www.tdw.com](http://www.tdw.com).

9 [www.funkanalyse.tns-infratest.com](http://www.funkanalyse.tns-infratest.com).

evaluation of specific programs. Depending on the specific *Land*, some of these surveys are replicated annually, others less often.

(8) As the MA only surveys a population older than fourteen years old, the KIM und JIM surveys, commissioned by the Research Consortium on Media Pedagogy (*Medienpädagogischer Forschungsverbund Südwest*)<sup>10</sup> specifically analyze media use by youth between the ages of six and thirteen, and between twelve and nineteen respectively. KIM started in 1999 and JIM in 1998; both are replicated annually and have sample sizes of about 1,200. In addition, for each child KIM surveys the legal parent or guardian (*Erziehungsberechtigte*) so that pairs of respondents can be analyzed. The topics of the KIM and JIM surveys include media use, leisure interests, sources of information, and – something unique for media studies – also television *preferences*. KIM additionally surveys media use in the context of the family.

The results from most of the above-mentioned studies are published in the monthly journal edited by *ARD-Werbung Sales & Services*, called *Media-Perspektiven*, which is considered even by the academic world to be one of “the three leading periodicals in communications science” (Hanitzsch and Altmeyen 2007). Thus, *Media-Perspektiven* continuously reports on the most recent trends in the use of periodicals (Vogel 2006), radio (Klingler and Müller 2008), and television (Zubayr and Gerhard 2008). Additionally, results are compiled in a yearly brochure, *Media-Perspektiven-Basisdaten*, the content of which can be found on the *Media Perspektiven* website, where this brochure can also be ordered.<sup>11</sup>

#### Surveys commissioned by the Federal Statistical Office and Academic Associations

Four sources respond to the questions discussed in section 2.2. All of them are multi-purpose surveys in the sense that they allow researchers to investigate attitudinal and behavioral correlates of media use. The first two allow for comparisons between European countries. (1) The 1999 Eurobarometer (EB), a survey commissioned by the EU every year, solicited information on media use from the then fifteen Member States. As these surveys also obtained information on environmental behavior and social capital (Schulz 2003; Wilke and Breßler 2005), the impact of media use on these domains could also be assessed. We strongly recommend to replicate the 1999 EB in future years. (2) The European Social Survey (ESS), which is financed by the European Science Foundation together with national funding agencies within individual European countries, has been administered since 2002 every second year in about twenty-five countries of East and West

---

<sup>10</sup> [www.mpfs.de](http://www.mpfs.de).

<sup>11</sup> <http://www.media-perspektiven.de/3921.html>.

Europe. Its core module, that is, the questions to be replicated in each wave, surveys the frequency of newspaper, radio, television, and Internet use for general and political purposes as well as voting behavior and some political, social, and religious attitudes. Thus, the ESS provides an opportunity to compare changes in media use in a macro perspective between European countries. More importantly, it allows researchers to assess the attitudinal and behavioral correlates of media use in the domains of politics, civil society, and religion.

Furthermore, two German sources provide time series on media use and attitudes to the media. (3) The General German Social Survey (ALLBUS, *Allgemeine Bevölkerungsumfrage der Sozialwissenschaften*) surveyed the use of and interest in many media in 1998 (Weiß 2001) and 2004. Furthermore, it surveyed trust in media in 1984, 1994, 2000, and 2002 (Schweiger 2007, 259). (4) The Time Budget Study of the Federal Statistical Office, administered in 1991 and 2001, describes the development of media use in the context of how leisure time has developed (Ehling 2004; Jäckel and Wollscheid 2004).

Both of these types of resources are underexploited today. However, it is generally the case that surveys commissioned by broadcasting agencies seem to contain a greater potential for analyzing media change, while the surveys commissioned by the Federal Statistical Office and academic associations provide more opportunities for examining attitudinal and behavioral correlates of media use.

## 4. Recommendations for the research infrastructure

### 4.1 Production of programs

- (1) *The establishment of a central media content archive.* In this archive, the contents of all media should be stored in digital form so that primary computer-aided content analyses (Maurer and Reinemann 2003, 62f) become feasible and available for secondary analyses. The following substantive orientations are proposed for this archive:
  - It should archive current productions, but simultaneously should also gather all materials already available in private or semi-private archives.
  - It should be more concerned with electronic media than with print media, given the current status of media content archiving.
  - It should archive video-type broadcasts as well as content analytic datasets constructed for their analysis – that is, code plans, results of

coder reliability tests with different category systems, and other materials required for replications and longitudinal analyses.

- It should instigate a call for content analytical longitudinal research projects, specifically in social problem areas where stand-alone studies up to now have prevailed. Moreover, it should support such projects while they are under way.

Formally, the archive should consist of at least two permanent positions devoted to data service funded by scientific agencies. It should rest on the open source principle, but it could – after it has been successfully established – require a fee for its services. Since up to now there have only been a few regional archives, such as the data archive of the Institute of Journalism (*Institut für Publizistik*) at the University of Mainz, these may form the core of the planned central archive.

In brief, to catch up with what has been accomplished in survey research since the 1950s, this report proposes the creation of a central infrastructure for content analyses – similar to the Leibniz Institute for the Social Sciences (GESIS, *Leibniz-Institut für Sozialwissenschaften*) but to serve the needs of survey research.

- (2) *Archiving content analytic time series of public agencies.* One of the tasks of the state media authorities for broadcasting in Germany (*Landesmedienanstalten*) and often also one of the voluntary endeavors of public broadcasting stations is to establish longitudinal data. They should be asked to extend their current research programs and to hand this over to archive the following data:
  - The biennial longitudinal content analyses of the radio program structure of the private radio stations financed by the media authorities for broadcasting of some of the German *Länder*. In the future, moreover, this research should be expanded to all the *Länder* and to public radio stations as well. Possibly, the ARD could take over part of the financing.
  - Thus far, the structure and quality of print media content have not been systematically evaluated. For this purpose, a research department should be established at the central institute.
- (3) *Archiving content analyses of individual researchers.* In contrast to public agencies, individual researchers are interested in specific theoretical questions rather than in long-term descriptions. Consequently, the content analyses of their stand-alone projects use different category systems. Nevertheless, these analyses should be gathered and prepared for secondary analyses in the archive as well. These systems are useful in the construction of more integrated and enduring category systems in future research.

- (4) *Developing a unified content analytical category system.* The German professional associations in the social and communication sciences should advocate the development of a system that contains the most general categories for the measurement of the program structure and quality in all media that establishes guidelines for more specific research projects. The research funding agencies should support this endeavor.
- (5) *Furthering cross-national comparisons.* The national professional societies in Europe should prepare a common core of content analytical categories. The European Science Foundation and individual national science funding agencies could finance pilot content analyses with the same category system in all European countries.

#### 4.2 Media consumption

- (1) *Providing privately funded data for the scientific community.* The German professional societies of the social and communication sciences should secure access to important surveys funded by media stations for scientific use. In particular, these are:
  - *AGF/GfK data.* As competing stations may understandably have some provisos against a premature release, a waiting period of some years should be contracted. Furthermore, as the AGF/GfK data are much richer and much more complicated than the MA data, a research project devoted to their transformation into meaningful indicators and, ultimately, time series that continue the ones constructed from the MA is recommended.
  - *the latest two editions of the Long-Term Study on Mass Communication (Langzeitstudie Massenkommunikation), 2000 and 2005.*
  - *Consumer studies, regional studies, and studies of specific audiences,* as mentioned in section 3.2.
- (2) *Enhancing the analytical potential of privately funded data.* Understandably, privately funded media surveys have been rarely concerned with more general social and political attitudes – the occasional questioning of party preferences in the MA being one of the rare exceptions. Yet, adding such questions can strongly enhance the public visibility of these studies and their funding agencies, once analyses are publicized. Therefore, the national professional associations should urge funders to include at least three more very brief, and therefore inexpensive, general questions that have been widely used in academic research: on party preference, church attendance, and union membership. If private funders cannot be motivated by appealing to their self-interest in terms of public

attention to future results, some financing through national academic funding agencies should be considered as well.

- (3) *Data linkage of programs and audience data.* Although the MA data, for example, contain exact information about the time someone uses a specific medium, content analysis and survey data are rarely linked, and programs and audiences are rarely analyzed simultaneously. Therefore, research projects that link content analyses and survey analyses should be supported.
- (4) *Feasibility project on comparisons of national data between countries.* Since there are consortia similar to the AG-MA in other countries that commission market research on the “advertising currency” and the quota of stations and broadcasting, a feasibility project that explores the possibility of comparisons should be launched. As national broadcasting stations will increasingly cooperate, their genuine commercial interests should be appealed to in order to support such an endeavor.
- (5) *Promoting cross-national comparisons.* Cross-national comparisons are rare in media research (Kleinstauber 2002: 56). Moreover, the cross-national comparison of media uses and media effects is on the bottom rather than on the top of the agenda of the German media research community (Wilke 2002: 18-31). There are two reasons for this
  - on the macro-level there is a lack of systematic research of media systems backed by quantitative indicators (Hallin and Mancini 2004). A European media indicator databank that collects indicators from the various sources mentioned (e.g., Thomaß 2007) would stimulate comparative research.
  - on the micro-level there is a lack of knowledge within the research community about existing cross-nationally comparative datasets (Livingstone 2003, Hanitzsch and Altmeppen 2007).

Therefore, the EU should be asked to replicate the 1999 EB on media use. Moreover, cross-national research using the above mentioned EB and ESS data should be encouraged.

## References:

- Amann, G. and Wipplinger, R. (1997): Sexueller Missbrauch in den Medien. In: Amann, G. and Wipplinger, R. (Eds.): Sexueller Missbrauch. Überblick zu Forschung, Beratung und Therapie. Ein Handbuch. Tübingen.
- Bauer, G. (Ed.) (2003): Programm-Profil-Analyse saarländischer Hörfunkprogramme 2003. Schriftenreihe der LMS Landesmedienanstalt Saarland, Band 10. Studie erstellt durch die Arbeitseinheit Medien- und Organisationspsychologie des Saarlandes. Projektleitung: Frank Schwab, Dagmar Unz.
- Bomas, W. (2005): Der duale Rundfunk. Sein Bedeutung für die Entwicklung des Rundfunkmarktes. Köln.
- Bonfadelli, H. (2002): Medieninhaltsforschung. Konstanz.
- Bonfadelli, H. (2007): Die Darstellung ethnischer Minderheiten in den Massenmedien. In: Bonfadelli, H. and Moser, H. (Eds.): Medien und Migration. Wiesbaden.
- Bonfadelli, H. and Moser, H. (Eds.) (2007): Medien und Migration. Wiesbaden.
- Bosch, E.M. (1988): Exkurs: Alter in der fiktiven Fernsehrealität – Eine Analyse der Konstruktion von Altersdarstellungen und ihrer Rezeption durch ältere Menschen. In: Eckhardt, J. and Horn, I. (Eds.): Ältere Menschen und Medien. Eine Studie der ARD/ZDF-Medienkommission. Frankfurt/M.
- Brosius, H.-B. and Rössler, P. (1999): Die soziale Realität in einfacher Pornographie und Softsex-Filmen. Ein Beitrag zur Pornographie-Diskussion. In: Rundfunk und Fernsehen 47 (1), 25-42.
- Brosius, H.-B. and Weiler, S. (2003): Programmanalyse nichtkommerzieller Lokalradios in Hessen. Eine Inhaltsanalyse im Auftrag der Hessischen Landesanstalt für privaten Rundfunk (LPR Hessen), München 2000, Schriftenreihe der LPR Hessen, Bd. 10.
- Brosius, H.-B. and Zubayr, C. (1996): Vielfalt im deutschen Fernsehprogramm. Eine empirische Anwendung eines Qualitätsmaßstabs. Rundfunk und Fernsehen 44 (2), 185 -214.
- Burgert, C. and Koch, T.K. (2008): Die Entdeckung der Neuen Alten? Best-Ager in der Werbung. In: Holtz-Bacha, C. (Ed.): Stereotype? Frauen und Männer in der Werbung. Wiesbaden.
- Davis, R. and Kubey, R. (1982): Growing Old on Television. In: Pearl, D./Bouthilet, L. and Lazar, J. (Eds.): Television and Behavior. Ten Years of Scientific Progress and Implications for the Eighties. Rockville.
- Donsbach, W. and Dupré, D. (1995): Mehr Vielfalt oder „more of the same“ durch mehr Kanäle? In: Mitchell, J. and Blumler, J.G. (Eds.): Fernsehen und Zuschauerinteressen. Baden-Baden.
- Ehling, M. (2004): Zeitbudgeterhebungen 1991/92 und 2001/2002 – Kontinuität und Wandel. In: Statistisches Bundesamt (Ed.): Alltag in Deutschland. Analysen zur Zeitverwendung. Wiesbaden.
- Engel, B. and Best, S. (2001): Mediennutzung und Medienbewertung im Kohortenvergleich. Media Perspektiven 11/2001, 554-563.
- Faul, E. (1988): Die Fernsehprogramme im dualen Rundfunksystem. Wissenschaftliche Begleitkommission zum Versuch mit Breitbandkabel in der Region Ludwigshafen/Vorderpfalz. 10. Band. Berlin.



- Faul, E. (1989): Die Fernsehprogramme im dualen Rundfunksystem. *Rundfunk und Fernsehen* 37 (1), 25-46.
- Fleming, K./Thorson, E. and Aitkin, C.L. (2004): Alcohol Adverting Exposure and Perceptions: Links with Alcohol Expectancies and Intentions to Drink or Drinking in Underaged Youth and Young Adults. *Journal of Health Communication* 9 (1), 3-30.
- Fröhlich, R. and Holtz-Bacha, C. (1995): *Frauen und Medien. Eine Synopse der deutschen Forschung*. Opladen.
- Fröhlich, R./Scherer, H and Scheufele, B. (2007): Kriegsberichterstattung in deutschen Qualitätszeitungen. *Publizistik* 52 (1), 11-32.
- Fürtjes, O. (2008): Die Entproletarisierung des Fußballpublikums im Kontext des gesellschaftlichen Wandels. Eine empirische Längsschnittanalyse zur Ursächlichkeit der schichtungshierarchischen Veränderung des Fußballpublikums, exemplifiziert an der Leserschaft des Kicker-Sportmagazins von 1954 bis 2005. Köln.
- Gerbner, G./Gross, L./Signorelli, N and Morgan, M. (1980): The „mainstreaming“ in America: Violence Profile No. 11. *Journal of Communication* 30 (3), 10-29.
- Gerbner, G./Gross, L./Signorielli, N./Morgan, M. and Jackson-Beeck, M. (1979): The demonstration of Power: Violence Profile N. 10. *Journal of Communication* 29 (3), 177-196.
- Giewald, O. and Heyen, A. (2007): *Lokales Fernsehen in Thüringen 2002*. <http://www.tlm.de>. [Last visited 11/11/2008].
- Gilles, D./Hagenah, J. and Meulemann, H. (2008): Freizeit zunehmend durch Fernsehen bestimmt. *Freizeit und Fernsehnutzung in Deutschland 1987-2005*. Informationsdienst Soziale Indikatoren 40, 11-14.
- Gleich, U. (1995): Hörfunkforschung in der Bundesrepublik. *Media Perspektiven* 11, 554-564.
- Gnädinger, C. (2007): Politikerinnen in deutschen Printmedien. Vorurteile und Klischees in der Berichterstattung. Saarbrücken.
- Goertz, L. (1996): Zwischen Nachrichtenverkündung und Infotainment. In: Hömberg, W. and Pürer, H. (Eds.): *Medien-Transformation*. Konstanz.
- Greger, V. (1998): Privatisierung politischer Berichterstattung im Fernsehen? In: Imhof, K. and Schulz, P. (Eds.): *Die Veröffentlichung des Privaten – Die Privatisierung des Öffentlichen*. Opladen.
- Groebel, J. and Gleich, U. (1993): *Gewaltprofil des deutschen Fernsehprogramms*. Schriftenreihe Medienforschung der Landesanstalt für Rundfunk Nordrhein-Westfalen; Band 6. Opladen.
- Hafez, K. (Ed.) (2002): *Die Zukunft der internationalen Kommunikationswissenschaft in Deutschland*. Hamburg.
- Hagemann, W. (1958): Der Wochenrhythmus der westdeutschen Tagespresse. *Publizistik* 3, 259-271.
- Hagenah, J. and Meulemann, H. (2009a): The Analytical Potentials of Survey Trend Data from Market Research. The Case of German Media Analysis Data. To appear. In: Baur, N. (Ed.): *Historische Sozialwissenschaft, Sonderheft*.
- Hagenah, J. and Meulemann, H. (2009b): Alte und neue Medien: Medientrends und sozialer Wandel. In: Hagenah, J. and Meulemann, H. (Eds.): *Alte und neue Medien – Zum Wandel der Medienpublika in Deutschland seit den 1950er Jahren*. Münster.

- Hallermann, K./Hufnagel, A./Schatz, K. and Schatz, R. (1998): Grundversorgung: Pflichten und Rechte. Eine Langzeit-Inhaltsanalyse zum Informationsangebot von ARD und ZDF. Bonn.
- Hallin, D.C. (1997): The Media and war. In: Corner, J./Schlesinger, P. and Silverstone, R. (Eds.): International Media Research. A Critical Survey. London/New York.
- Hallin, D.C. and Mancini, P. (2004): Comparing Media Systems. Three Models of Media and Politics. Cambridge.
- Handel, U. (1998): Tatort Hannover. Ausländer und Deutsche in der Kriminalitätsberichterstattung im Vergleich. In: Quandt, S. and Gast, W. (Eds.): Deutschland im Dialog der Kulturen: Medien, Images, Verständigung. Konstanz.
- Hanitzsch, T. and Altmeyden, K.-D. (2007): Über das Vergleichen: Komparative Forschung in den deutschen kommunikationswissenschaftlichen Fachzeitschriften. *Medien & Kommunikationswissenschaft* 55 (2), 185-203.
- Hasebrink, U. (2006): Hörfunklandschaft in Schleswig-Holstein. Bestandsaufnahme 20 Jahre nach Einführung der dualen Rundfunkordnung. Kiel.
- Held, B. and Simeon, Th. (1994): Die zweite Stunde Null. Berliner Tageszeitungen nach der Wende (1989-1994). Marktstrukturen, Verlagsstrategien, publizistische Leistungen. Berlin.
- Hesse, M. (2001): Wer macht die Nachrichten. Über die öffentliche Sichtbarkeit von Männern und Frauen in den Medien 2000. Forschungsbericht Nr. 105. Bonn.
- Heyen, A. (2001): Programmstrukturen und Informationsangebote im Radio. Ergebnisse und Erfahrungen aus sechs Jahren Programmforschung der TLM. In: Rössler, P./Vowe, G. and Henle, V. (Eds.): Das Geräusch der Provinz – Radio in der Region. München.
- Hüther, J./Scholand, H. and Schwarte, N. (1973): Inhalt und Struktur regionaler Großzeitungen. Düsseldorf.
- Jäckel, M. and Wollscheid, S. (2004): Mediennutzung im Tagesverlauf: Ausweitung des Angebots und Strukturen der Zeitverwendung. In: Statistisches Bundesamt (Ed.): Alltag in Deutschland. Analysen zur Zeitverwendung. Wiesbaden.
- Jud-Krepper, H. (1997): Das Mädchen- und Jungenbild im Fernsehen und seine sozialisierende Wirkung. In: Erlinger, H.-D. (Ed.): Neue Medien, Edutainment, Medienkompetenz. Deutschunterricht im Wandel. München.
- Kaase, M. (1989): Fernsehen, gesellschaftlicher Wandel und politischer Prozeß. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, Sonderheft 30, 97-117.
- Kleinsteuber, H.J. (2002): Mediensysteme im internationalen Vergleich: Ein Überblick. In: Hafez, K. (Ed.): Die Zukunft der internationalen Kommunikationswissenschaft in Deutschland. Hamburg.
- Kleinsteuber, H.J. and Wiesner, V. (1992): Szenario des Niedergangs. Die Entwicklung des öffentlichen Rundfunks in Australien und Kanada als Modell für Europa? *Medium* 22 (1), 57-60.
- Kleinsteuber, H.J./Wiesner, V. and Wilke, P. (1991): Public Broadcasting im internationalen Vergleich. Analyse des gegenwärtigen Stands und Szenarien einer zukünftigen Entwicklung. *Rundfunk und Fernsehen* 29 (1), 33-54.
- Klingler, W and Müller, D.K. (2008): Radio II: Stabile Nutzungsmuster auch bei erweiterter Grundgesamtheit. Radionutzung erstmals inklusive EU-Ausländer und Kinder ab zehn Jahren. *Media Perspektiven* 10/2008, 502-515.

- Kolmer, C. (2004): Krieg im Fernsehen 2001-2003. Eine Inhaltsanalyse der deutschen, englischen, südafrikanischen und US-Fernsehnachrichten. In: *Zeitschrift für Kommunikationsökologie* 6 (1), 37-43.
- Krüger, U.M. (1992): *Programmprofile im dualen Fernsehsystem 1985-1990*. Baden-Baden.
- Krüger, U.M. (2001): *Programmprofile im dualen Fernsehsystem 1991-2000*. Baden-Baden.
- Krüger, U.M. (2005a): Themenprofile deutscher Fernsehnachrichten. In: *Media Perspektiven* 7/2005, 302-320.
- Krüger, U.M. (2005b): Sparten, Sendungsformen und Inhalte im deutschen Fernsehangebot. *Media Perspektiven* 5/2005, 190-205.
- Krüger, U.M. (2008): InfoMonitor 2007: Unterschiedliche Nachrichtenkonzepte bei ARD, ZDF, RTL und Sat.1. *Media Perspektiven* 2/2008, 58-83.
- Krüger, U.M. and Zapf-Schramm, T. (2002): Öffentlich-rechtliches und privates Fernsehen: Typische Unterschiede bleiben bestehen. *Media Perspektiven* 4/2002, 178-190.
- Krüger, U.M. and Zapf-Schramm, T. (2008): Sparten, Sendungsformen und Inhalte im deutschen Fernsehangebot 2007. Programmanalyse von ARD/Das Erste, ZDF, RTL, SAT.1 und ProSieben. *Media Perspektiven* 4/2008, 166-189.
- Küchenhoff, E. (1975): *Die Darstellung der Frau und die Behandlung von Frauenfragen im Fernsehen: eine empirische Untersuchung*. Stuttgart et al.
- Kunczik, M. (2008): *Medien und Gewalt: Der aktuelle Forschungsstand*. In: Dittler, U. and Hoyer, M. (Eds.): *Aufwachsen in virtuellen Medienwelten: Chancen und Gefahren digitaler Medien aus medienpsychologischer und medienpädagogischer Perspektive*. München.
- Lindner-Braun, C. (2007): *Mediennutzung. Methodologische, methodische und theoretische Grundlagen*. Berlin.
- Livingstone, S. (2003): On the challenges of Cross-national Comparative Media Research. *European Journal of Communication* 18 (4), 477-500.
- Maier, M. (2002): *Zur Konvergenz des Fernsehens in Deutschland. Ergebnisse qualitativer und repräsentativer Zuschauerbefragungen*. Konstanz.
- Maier, M./Ruhrmann, G. and Klietsch, K. (2006): *Der Wert von Nachrichten im deutschen Fernsehen. Ergebnisse einer Inhaltsanalyse 1992-2004*. Düsseldorf. <http://www.lfm-nrw.de>. [Last visited 01/18/2009]
- Marchal, P. (2004): *Kultur und Programmgeschichte des öffentlich-rechtlichen Hörfunks in der Bundesrepublik Deutschland. Band II: Von den 60er Jahren bis zur Gegenwart*. München.
- Maurer, M. and Reinemann, C. (2003): *Schröder gegen Stoiber – Nutzung, Wahrnehmung und Wirkung der TV-Duelle*. Wiesbaden.
- Maurer, M. and Reinemann, C. (2006): *Medieninhalte. Eine Einführung*. Wiesbaden.
- mediatenor.de (2009): *Media Tenor – Medien-Monitoring & Medieninhaltsanalyse*. <http://www.mediatenor.de>. [Last visited 01/18/2009].
- Meier, H.E. (2003): Beyond Convergence. *European Journal of Communication* 18 (3), 337-365.
- Meier, K. (2002): *Ressort, Sparte, Team. Wahrnehmungsstrukturen und Redaktionsorganisation im Zeitungsjournalismus*. Konstanz.
- Merten, K. (1994): *Konvergenz der deutschen Fernsehprogramme. Eine Langzeituntersuchung 1980-1993*. Münster.

- Merten, K. (1996): Konvergenz der Fernsehprogramme im dualen Rundfunk. In: Hömberg, W. and Pürer, H. (Eds.): Medien-Transformation. Konstanz.
- Meulemann, H. (2009): Verdrängt das Internet die Massenmedien? Medienpräferenzen und die Individualisierung der Mediennutzung. Schriftenreihe des Instituts für Rundfunkökonomie an der Universität zu Köln.
- Meulemann, H./Hagenah, J. and Gilles, D. (2009): Neue Angebote und alte Gewohnheiten. Wie und warum das deutsche Publikum zwischen 1987 und 1996 vom öffentlich-rechtlichen auf das private Fernsehen gewechselt hat. Publizistik 54 (2), 240-264.
- Meyen, M. (2004): Mediennutzung. 2<sup>nd</sup> edition. Konstanz.
- Moses, K. and Heyen, A. (2003): Lokales Fernsehen in Thüringen 2002. <http://www.tlm.de>. [Last visited 12/12/2008].
- Olien, C./Tichenor, Ph. and Donohue, G. (1989): Media Coverage and Social Movements. In: Salmon, C. (Ed.): Information Campaigns. Balacing Social Values and Social Change. Newbury Park/London/New Delhi.
- Peiser, W. (1996): Die Fernsehgeneration. Eine empirische Untersuchung ihrer Mediennutzung und Medienbewertung. Opladen.
- Petersen, J. (2006): Männlichkeit. Eine Inhaltsanalyse ihrer Darstellung in Männer- und Frauenzeitschriften in Deutschland. Berlin.
- Petersen, T. and Jandura, O. (2006): Testing visual signals in representative surveys in combination with media content analyses of the 2002 German Federal election campaign. International Journal of Public Opinion Research 19 (1), 89-96.
- Petzold, T. (2008): Gewalt in internationalen Fernsehnachrichten: Eine komparative Darstellung medialer Gewaltpräsentation in Deutschland, Großbritannien und Russland. Wiesbaden.
- Pfetsch, B. (1991): Die Fernsehformate von Politik im Dualen Rundfunksystem. In: Gellner, W. (Ed.): An der Schwelle zu einer neuen deutschen Rundfunkordnung. Berlin.
- PriceWaterhouseCoopers (2007): Report for BBC Trust Unit. BBC On-Demand Proposals – Assessment of BBC Managements's "New Media" Assumptions and Assertions. [http://www.bbc.co.uk/bbctrust/assets/files/pdf/review\\_report\\_research/pvt\\_iplayer/bbctrust\\_pwcreport\\_pva\\_annexe.pdf](http://www.bbc.co.uk/bbctrust/assets/files/pdf/review_report_research/pvt_iplayer/bbctrust_pwcreport_pva_annexe.pdf). [Last visited 12/12/2008].
- Rager, G. and Siebers, T. (2006): Inhaltsanalyse des Senders RPR 1. Ludwigshafen. [Unpublished research report].
- Rager, G./Siebers, T. and Hassemer, G. (2005): Hörfunk 2005 in Mecklenburg-Vorpommern. Programmanalyse, Strukturen und Potenziale. Schwerin.
- Reitze, H. and Ridder, C.-M. (2006): Massenkommunikation VII – eine Langzeitstudie zur Mediennutzung und Medienbewertung 1964-2005. Baden-Baden.
- Rose, C. (1995): Die Drogenberichterstattung der deutschen überregionalen Tagespresse. Ergebnisse einer Inhaltsanalyse von FAZ, FR, taz und Welt. Sucht 41 (1), 34-42.
- Rossen-Stadtfeld, H. (2005): Funktions und Bedeutung des öffentlich-rechtlichen Kulturauftrags im dualen Rundfunksystem. Köln.
- Rossmann, C./Brandl, A. and Brosius, H.-B. (2003): Der Vielfalt eine zweite Chance? Publizistik 48 (4), 427-454.

- Ruhrmann, G. (2002): Wie regionale Tageszeitungen über Migranten berichten. Das Beispiel Thüringen. In: Meier-Braun, K.-H. and Kilgus, M.A. (Eds.): *Integration durch Politik und Medien?* Baden-Baden.
- Ruhrmann, G. and Kollmer, J. (1987): *Ausländerberichterstattung in der Kommune. Inhaltsanalyse Bielefelder Tageszeitungen unter besonderer Berücksichtigung ausländerfeindlicher Alltagstheorien.* Opladen.
- Saxer, U. (1980): Führt ein Mehrangebot an Programmen zu selektivem Rezipientenverhalten? *Media Perspektiven* 6/1980, 395-406.
- Schatz, H./Immer, N. and Marcinowski, F. (1989): Der Vielfalt eine Chance? Empirische Befunde zu einem zentralen Argument für die Dualisierung des Rundfunks in der Bundesrepublik Deutschland. *Rundfunk und Fernsehen* 37 (1), 5-24.
- Scheufele, B. (2005): *Sexueller Missbrauch – Mediendarstellungen und Medienwirkung.* Wiesbaden.
- Schmerl, C. (Ed.) (1985): *In die Presse geraten: Darstellung von Frauen in der Presse und Frauenarbeit in den Medien.* Köln.
- Scholz, R. (1993): *Konvergenz im TV-Sport. Eine komparative Studie des „Dualen Fernsehsystem“.* Berlin.
- Schönbach, K./Feierabend, S. and Möhring, W. (1993): *Lokale Hörfunksender in Baden-Württemberg. Eine Inhaltsanalyse ihrer Berichterstattung, LFK-Schriftenreihe Dialog, Bd. 7.* Stuttgart.
- Schulz, W. (1970): *Der Inhalt der Zeitungen. Eine Inhaltsanalyse der Tagespresse in der Bundesrepublik Deutschland (1967) mit Quellentexten früher Inhaltsanalysen Amerika, Frankreich und Deutschland.* Düsseldorf.
- Schulz, W. (2003): Mediennutzung und Umweltbewusstsein: Dependenz- und Priming-Effekte. *Publizistik* 48 (4), 387-483.
- Schulz, W. (2003): *Politische Kommunikation.* In: Bentele, G./Brosius, H.B. and Jarren, O. (Eds.): *Öffentliche Kommunikation.* Wiesbaden.
- Schwantag, K. (1974): *Die betriebliche Anpassung lokaler und regionaler Abonnementszeitungen an die durch intra- und intermediären Wettbewerb der Massenkommunikationsmittel ausgelösten Veränderungen der Leserbedürfnisse.* In: Presse- und Informationsamt der Bundesregierung (Ed.): *Kommunikationspolitische und kommunikationswissenschaftliche Forschungsprojekte der Bundesregierung (1971-1974).* Bonn.
- Schweiger, W. (2007): *Theorien der Mediennutzung. Eine Einführung.* Wiesbaden.
- Semetko, H.A. (2009): Election campaigns, partisan balance, and the news media. In: Norris, P. (Ed.): *The Roles of the News Media: Watch-dogs, Agenda-setter and Gate-Keepers.* Washington.
- Stein-Hilbers, M. (1977): *Kriminalität im Fernsehen. Eine inhaltsanalytische Untersuchung. Kriminalität und ihre Verwalter, Bd. 6.* Stuttgart.
- Stuiber, H.-W. (1990): *Landesweiter Hörfunk in Bayern – Programm, Publikums-wünsche und Bewertungen: Inhaltsanalyse, Image- und Akzeptanzstudie zu den Hörfunkprogrammen Antenne Bayern, Bayern 1 und Bayern 3.* München.
- Sutor, S. (1999): *Programmbeobachtung der BLM.* In: Jarren, O. (Ed.): *Rundfunk in öffentlich-rechtlicher Trägerschaft. Modell für modernes Rundfunkmanagement / BLM Symposium Medienrecht 1999.* München.
- Thomaß, B. (Ed.) (2007): *Mediensysteme im internationalen Vergleich.* Konstanz.
- t1m.de (2000): *Lokales Fernsehen in Thüringen 1999.* <http://www.t1m.de>. [Last visited 12/23/2008].

- Trebbe, J. (2004): Fernsehen in Deutschland 2003-2004. Programmstrukturen, Programminhalte, Programmentwicklungen. Berlin.
- Trebbe, J. and Weiß, H.-J. (1994): Öffentliche Streifragen in privaten Fernsehprogrammen. Opladen.
- Ulze, H. (1977): Frauenzeitschrift und Frauenrolle: eine aussagenanalytische Untersuchung der Frauenzeitschriften Brigitte, Freundin, Für Sie und Petra. Berlin.
- Vogel, A. (2006): Stagnation auf hohem Niveau. Media Perspektiven 7/2006, 380-398.
- Volpers, H. (2009): Hörfunklandschaft Niedersachsen 2009. Projektbeginn: Januar 2009, Auftraggeber: Niedersächsische Landesmedienanstalt (NLM). <http://www.alm.de>. [Last visited 02/09/2009].
- Volpers, H. and Schnier, D. (1996): Das WDR-Hörfunkprogramm Eins Live. Ergebnisse einer empirischen Programmanalyse. Media Perspektiven 5/1996, 249-258.
- Volpers, H./Salwiczek, C. and Schier, D. (2008): Die Regionalfenster von RTL und Sat.1 im Jahr 2006. Ergebnisse einer kontinuierlichen Programmanalyse. ALM Programmbericht 2007.
- Weiderer, M. (1992): Das Frauen und Männerbild im deutschen Fernsehen. Eine inhaltsanalytische Untersuchung der Programme von ARD, ZDF und RTL plus. Regensburg.
- Weiß, H.-J. (1999): Programmalltag in Deutschland. In: Arbeitsgemeinschaft der Landesmedienanstalten in Deutschland (Ed.): Programmbericht zur Lage und Entwicklung des Fernsehens in Deutschland 1998/99. Berlin.
- Weiß, H.-J. (2007): Private Fernsehvollprogramme 1998-2007. Eine 10-Jahres-Bilanz der kontinuierlichen Fernsehprogrammforshung der Landesmedienanstalten. In: ALM Programmbericht 2007.
- Weiß, H.-J. and Trebbe, J. (2000): Fernsehen in Deutschland 1998-1999. Programmstrukturen – Programminhalte – Programmentwicklungen. Schriftenreihe der Landesmedienanstalten, Bd. 18.
- Weiß, R. (2001): Der praktische Sinn des Mediengebrauchs im Alltag. In: Maier-Rabler, U. and Latzer, M. (Eds.): Kommunikationskulturen zwischen Kontinuität und Wandel. Konstanz.
- Werner, P. and Rinsdorf, L. (1998): Ausgeblendet? Frauenbild und Frauenthemen im nordrhein-westfälischen Lokalfunk. Opladen.
- Wichert, L. (1997): Radioprofile in Berlin-Brandenburg 1996. Zehn privaten und ein öffentlich-rechtliches Programm im Vergleich. Schriftenreihe der mabb, Band 6.
- Wichert, L. (2008): Radioprofile in Brandenburg. Fünf Lokalprogramme im Vergleich 2007. Schriftenreihe der mabb, Band 23.
- Wilke, J. (2002): Internationale Kommunikationsforschung. Entwicklungen, Forschungsfelder, Perspektiven. In: Hafez, K. (Ed.): Die Zukunft der internationalen Kommunikationswissenschaft in Deutschland. Hamburg.
- Wilke, J. and Breßler, E. (2005): Europa auf dem Weg in die Informationsgesellschaft? Eine Auswertung von Eurobarometer-Daten. In: Rössler, P. and Krotz, F. (Eds.): Mythen der Mediengesellschaft – The Media Society and its Myths. Konstanz.
- Zubayr, C. and Gerhard, H. (2008): Tendenzen im Zuschauerverhalten, Fernsehgewohnheiten im Jahr 2007. Media Perspektiven 3/2008, 106-119.

## 6.5 Judicature

*Wolfgang Heinz*

Contact:

Wolfgang Heinz  
University of Constance  
Faculty of Law  
Universitätsstraße 10 · Box D 119  
78457 Konstanz  
Germany  
e-mail: wolfgang.heinz[at]uni-konstanz.de



## Abstract

The German crime statistics system should provide an adequate database for addressing questions about the level, structure, and variation of crime rates; prosecution and sanctioning by the authorities; the level and types of criminal sanctions imposed; and the enforcement of sanctions and reconviction rates. In these areas, the German criminal statistics system suffers many deficiencies. It is therefore necessary to improve the existing statistics by carrying out periodic crime and victimization surveys, surveys of suspects in the preliminary stages of public prosecution, and by compiling statistics on the enforcement of criminal offenses and reconviction rates.

But comprehensively optimizing the criminal statistics system would require a statistical database that contains all the data from police crime statistics and all judicial decisions relevant for criminological research, and then linking these data with pseudonymized individual data. This statistical database would solve the problems of the existing German crime statistics and would offer a basis for new, regularly compiled federal statistics, in particular on the execution of sentences and recidivism, as well as for caseflow statistics and cohort studies.

Keywords: caseflow statistics, cohort studies, conviction statistics, crime, sanctions, criminal, police crime statistics, prison statistics, probation service statistics, recidivism, reconviction statistic, statistics of criminal courts, statistics of the public prosecution offices, victimization survey

## 1. The existing criminal statistics system

The criminal statistics system in Germany consists of statistics that cover the areas of the police, public prosecutors, criminal courts, probation services, and prisons. The data is collected at the state level, where it is checked for plausibility and published in the statistics of the *Länder*. The Federal Criminal Police Office (BKA, *Bundeskriminalamt*) and the Federal Statistical Office summarize the aggregate data from the *Länder* statistics (see figure 1).

The statistics compiled by the public prosecution offices and criminal courts are procedure statistics (the statistical units are procedures), while the other statistics mentioned in figure 1 – Police Crime Statistics, Conviction Statistics, Probations Service Statistics, Prison Statistics – are personal statistics (the statistical units are individuals). In the personal statistics, the statistical units always designate the sex of the individual as well as their age

Figure 1: Overview of the crime and criminal justice statistics in the Federal Republic of Germany

Stage of Procedure (statistical unit)	Data collection (statistics published at the federal level)
<b>Preliminary proceedings</b>	
<b>Police Investigation</b> (Criminal suspicion: case, criminal suspect, victim)	<b>Police Crime Statistics (<i>Polizeiliche Kriminalstatistik</i>)</b> (Federal Criminal Police Office) (since 1953)
<b>Public prosecutor's decision on the result of the investigation</b> (Cases dealt with by the public prosecution office. Statistical unit is procedure; since 1998 the statistics also contain information on individuals)	<b>Statistics of the public prosecution offices (StAStat, <i>Staatsanwaltschaftsstatistik</i>)</b> (Federal Statistical Office) (since 1981)
<b>Main proceedings</b>	
<b>Proceedings of the criminal courts</b> (Criminal prosecutions processed by the courts. Statistical unit is procedure; since 1989 the statistics inform also on persons)	<b>Statistics of criminal courts (<i>Justizgeschäftsstatistik in Strafsachen</i>)</b> (Federal Statistical Office) (since 1959)
<b>Decisions of the criminal courts</b> (sentencing, conviction, based on persons)	<b>Conviction Statistics (<i>Strafverfolgungsstatistik</i>)</b> (Federal Statistical Office) (since 1950)
<b>Conviction / prison</b>	
<b>Suspension of sentence for parole</b> (Placed under the supervision of a full-time probation officer) (Waiver/revocation of parole is based on the subject)	<b>Probation Service Statistics (<i>Bewährungshilfestatistik</i>)</b> (Federal Statistical Office) (since 1963)
<b>Execution of a prison sentence</b> (Number and type of prisons, actual population, capacity, demographic characteristics of the prisoners)	<b>Prison Statistics (<i>Strafvollzugsstatistik</i>)</b> (Federal Statistical Office) (since 1961)

(in age groups<sup>1</sup>). These statistics also contain the citizenship of the person, which in the Police Crime Statistics (PCS, *Polizeiliche Kriminalstatistik*) is determined by using a comprehensive citizenship key (*Staatsangehörigkeits-schlüssel*) covering all nationalities. Currently, only the citizenship of the main guest worker countries is accounted for in the criminal justice statistics (this will be changed in 2009; see section 3.2.2 below). For the particularities of the statistical units and statistical variables, refer to the overview in the First Periodical Security Report (Federal Ministry of the Interior and Ministry of Justice 2001b: 15; see also Heinz 2003; 2008).

## **2. Research questions**

The central issues that should be answered by the criminal statistics system are questions about the level, structure, and variation of crime rates; prosecution and sanctioning by the authorities; the level and types of criminal sanctions imposed; and the enforcement of sanctions and reconviction rates.

## **3. Status quo of the crime and criminal justice statistics: databases and access**

### *3.1 Information deficits in the existing criminal statistics system*

When measured against the aforementioned research questions, the existing criminal statistics system in Germany has many shortcomings, despite the fact that it has been progressively expanded to include new statistics in various sectors (see figure 1) and additional or more differentiated statistical variables. Yet even with these improvements, the following fundamental problems have not been overcome:

- The existing criminal statistics system is limited to officially reported crime.

---

1 In the Police Crime Statistics (PCS), children are assigned to the following age groups: under 6, 6–8, 8–10, 10–12, and 12–14. For persons above the age of criminal responsibility, the PCS, the Conviction Statistics, and the Probation Service Statistics contain the following age groups: youth (in the PCS and in the Probation Service Statistics 14–16, 16–18), adolescents (18–21), young adults (in the PCS 21–23, 23–25), 25–30, 30–40, 40–50, 50–60, 60 and older (in the Conviction Statistics 60–70, 70 and older). In the Prison Statistics the age groups are even more highly differentiated in some areas.

- The current crime and criminal justice statistics are not coordinated either in content or methodology. Statistical units and variables are only partially compatible. The statistics are compiled according to different measurement and processing standards; an input-output analysis of the system is therefore not possible. Additional data collection beyond the stipulated statistical variables is a rare exception. The lack of compatibility is not just a problem between the PCS and the criminal justice statistics, but also among the different sets of criminal justice system statistics.
- The process of prosecution and sanctioning cannot be examined for specific offense groups. To do so one would need to be able to link data collected at the individual level to different sets of statistics (caseflow statistics), which is currently impossible. This in turn would require the transmission and storage of pseudonymized individual data.
- Data collection is carried out by authorities of the *Länder*. Checking the validity of the data in terms of both content and adherence to formal requirements is only possible at the state level; at the federal level it is only possible to conduct plausibility and consistency tests.
- Until very recently, Criminal Police Offices and Statistical Offices of the *Länder* did not provide federal statistical authorities (BKA, Federal Statistical Office) with any individual statistical data (microdata) (see section 3.2 below for recent changes).
- Regional criminal justice statistics are partially incomplete. In some cases they are not reported at all. For example, conviction statistics were not compiled in Saxony-Anhalt until 2007; probation service statistics are not being compiled in three out of five of the new *Länder*; and statistics on persons committed by a criminal court to a psychiatric hospital or institution for drug rehabilitation are not being compiled in four out of five of the new *Länder*. In other cases, compilation of the statistics has been suspended (e.g., the probation service statistics in Hamburg since 1997 and in Schleswig-Holstein between 2002 and 2006). This is due to the fact that there is no federal law that compels the *Länder* to collect the data for criminal justice statistics.
- The PCS cover only a portion of the actual registered crime rates; road traffic crime rates, for example, are left out. Due to a lack of adequately differentiated statistical variables in the criminal justice statistics, the statistics do not accurately reflect the decisions of the public prosecutors or the criminal convictions or executions of sanctions.
- Given the statistical units and variables as well as the measurement and processing standards currently available, the existing criminal statistics

provide a very inadequate basis for fundamental scientific research. This is due in part to the fact that the crime and criminal justice statistics are collected and processed according to regulations stipulated by federal and state government committees in which there is no one representing the research community.

- Scientific research is most productive when it has access to individual data records rather than aggregate data for statistical purposes. At present, only the individual data records of some of the criminal justice statistics are available from Research Data Centers of the Federal Statistical Office and the Statistical Offices of the German *Länder*. The individual data records of the PCS are not available through the Research Data Centers.

The 2001 KVI report already listed several of these deficits and emphasized the need for improvement by collecting caseflow statistics and additional information. These improved statistics would better meet the research demand for high-quality data linked to individual data records at the state level and the need for statistics on the enforcement of criminal sanctions such as fines or educative and disciplinary measures under juvenile law (see Albrecht 2001: 66f).

### *3.2 Changes in databases and data access since the 2001 KVI report*

#### *3.2.1 Extension of the criminal statistical systems on unreported crimes, on fear of crime and on the subjective recognition of crime rates and crime control*

In its First Periodical Report on Crime and Crime Control, the German Federal Government stated that continually updated research in the area of unreported crime is “a necessary instrument for measuring developments in crime for those types of offense where this is appropriate” (Federal Ministry of the Interior and Federal Ministry of Justice 2001b: 600). In early 2002, a working group (AG BUKS) was set up by the Ministry of the Interior and the Ministry of Justice with the task of creating a concept for a periodical crime survey in order to gain insight into the extent of victimization, the scale of reporting behavior and experiences of crime victims when reporting to the police, and public attitudes on various aspects of crime and the fear of crime. The final report of this working group was submitted to the commissioning ministries in September 2002 (Heinz 2002). The recommendations have not yet been implemented, mainly due to the costs involved.

### 3.2.2 Changes in databases and data access in the field of crime and criminal justice statistics

#### 3.2.2.1 Databases

*Police Crime Statistics (PCS).* In 1997/98 the police internally developed a plan to install “PCS-new” as a component of a comprehensive executive information system, which was abandoned in 2002 due to serious difficulties in implementation. From this date on, the further development of PCS-new has been carried out solely in the framework of the police information system INPOL (*Informationssystem der Polizei*) analysis. At the end of 2004, the “old” database of the PCS was integrated to the new Oracle database of the INPOL systems. PCS-new will be introduced in a two-stage procedure. Basically, inputting of the individual data records to the current PCS and conversion from a four-digit to a six-digit criminal offense code will take place in the first stage. The second stage will include expanding the PCS system to incorporate additional catalogs that allow more differentiated collection of individual variables.

Since 1 January 2007, the BKA has been collecting pseudonymized individual data records. Since not all of the *Länder* were in a position to provide this data at the outset, a transitional period was established. During this period, individual data records were provided to the BKA along with the aggregated tables. It was planned that by 2008, the delivery of individual data records would be carried out nationwide, at which point aggregated delivery would be discontinued. Ideally, the BKA should currently be in a position to compile statistics for each of the *Länder* as well as for the federal government on the basis of these individual data records.

The pseudonymized individual data records allow personal data to be collated in order to determine the “real” number of suspects at the federal level. Furthermore, by providing access to the individual data records, the BKA makes it possible for the first time to link the variables collected and to use different methods of statistical analysis.

The previous four-digit criminal offense code will be converted to a six-digit code. This will lay the foundation for a far more complete and differentiated record of criminal offenses than under previous standards. It will increase the currently 421 code numbers to approximately 1,500 code numbers.

*Statistics of the public prosecutors offices and criminal courts.* Since the reporting year 2004, the Federal Statistical Office provides (anonymized) individual data from the *Länder* to the public prosecutors offices and criminal courts for use in their statistics. This enables the Federal Statistical Office to flexibly analyze data available beyond the regularly published statistics on public prosecutors and criminal courts whenever a need arises.

Since 2004, criminal proceedings initiated by public prosecutors have been categorized following a subject matter catalog of the penal provisions violated, in which there are currently 30 categories. Detailed information is published on only six of these categories.

*Conviction Statistics, Prison Statistics, Probation Service Statistics.* For 2007, the specific offense codes used in each set of criminal justice statistics were replaced by a uniform recommended code system for all personal statistics. This will enable better comparison of the results of the three statistics in regard to convictions for felonies. This will constitute a major advance in the field of personal criminal justice statistics.

From 2009 on, the conviction and prison statistics were to have removed the previous restriction on data collection concerning citizenship, which had been differentiated only according to either German or non-German citizenship as well as citizenship of a few important guest worker groups. In the future, citizenship of convicted criminals or prisoners and detainees will be contained with the complete area and citizenship code in the official population statistics. In contrast, probation service statistics will continue to differentiate only between Germans and non-Germans.

*Conviction Statistics.* Since 2007, conviction statistics have been compiled in Saxony-Anhalt, making it possible to now publish statistics covering all of Germany. From the reporting year 2009 forward, the levels of criminal sanctions currently recorded in fixed categories will be recorded in a non-categorized manner.

#### 3.2.2.2 Data access by data users

Since 1997, the annual report of the PCS has also been published on the Internet; the annual tables and time series dating back to 1987 can be downloaded as PDF files;<sup>2</sup> interested users can also receive this data as an Excel file. Up to the end of the 1990s, the Federal Statistical Office published the technical series *Strafrechtspflegestatistiken* exclusively in printed format, and for a few years in both printed and electronic form. Since 2004, only the electronic version has been provided<sup>3</sup> (Excel or PDF). Time series on convicted criminal offenders as well as convicted Germans and foreigners are also available for downloading.

Since autumn 2005, microdata from the Conviction Statistics (StVerfStat), Prison Statistics (StVollzStat), and since summer 2008 also from the Probation Service Statistics (BewHiStat) on the reporting years since 1995 have been made available for scientific analysis through the Research Data Centers of the Federal Statistical Office and the Statistical Offices of the

---

2 <http://www.bka.de/>, click on Reports and Statistics, Crime Statistics.

3 <https://www-ec.destatis.de>

German *Länder*.<sup>4</sup> At present the data in the judicial statistics are solely available through controlled remote data processing.<sup>5</sup> In addition, what is known as a “Public Use File” is being prepared for the Conviction Statistics.

The annual reports of the PCS contain not only data in tables, but also detailed explanatory reports, time series, graphic visualizations, and interpretations.

On the other hand, the Federal Statistical Office limits itself primarily to the publication of the criminal justice statistics in tabular form, which are supplemented by several datasets with time series. There are, however, noteworthy exceptions to this publication system. In 1999, the results of the various criminal justice statistics were summarized by the Federal Statistical Office in the publication *Justice as Reflected in the Criminal Justice Statistics*. The brochure *Justice at A Glance*,<sup>6</sup> published in 2008, included data provided by the statistical offices as well as results of scientific research.

For the first time in 2001 and again in 2006, the federal government sought to present a comprehensive picture of the crime situation in Germany in its two *Periodical Reports on Crime and Crime Control in Germany* (Federal Ministry of the Interior and Federal Ministry of Justice 2001; 2006). The aim was

“to put together the most comprehensive picture possible of the crime situation in Germany. For the first time in a report, it will draw together findings taken from the existing pool of official data [...]. At the same time, the report will provide a useful reference on the results of scientific research into the incidence and causes of crime” (Federal Ministry of the Interior and Federal Ministry of Justice 2001a: 3).

Through the inclusion of data from other data sources and the results of scientific research (e.g., victim surveys), the two reports provided a synopsis of the state of knowledge in the field of crime and internal security.

---

4 <http://www.forschungsdatenzentrum.de/en/>

5 On remote data processing, or “remote execution,” see <http://www.forschungsdatenzentrum.de/en/datenzugang.asp>

6 <http://www.destatis.de/jetspeed/portal/cms/Sites/destatis/Internet/DE/Content/Publikationen/Broschueren/JustizBlick,property=file.pdf>



#### **4. Optimizing the existing criminal statistics systems: Recommendations of the German Data Forum (RatSWD) working group**

##### *4.1 Tasks of the working group*

On the 27 October 2006, the German Data Forum (RatSWD) hosted the workshop “Data Problems in the Crime and Criminal Justice Statistics” in Berlin. Participants at this workshop adopted a memorandum in which the RatSWD was asked to assemble a working group tasked to create “proposals for a comprehensive optimization of the existing criminal statistics systems.”<sup>7</sup> In response to the memorandum, the RatSWD assembled a working group of eleven members. The task of the working group was to deliver proposals for the “optimization of the existing criminal statistics systems” within a year. The constitutive meeting of the group took place in Berlin in July 2007.

With regard to the aforementioned deficits listed in the 2001 KVI report and additional information gaps (see 3.1 above) that the working group had noted as deficiencies, proposals were made to redress these deficiencies in the short, medium, and long term. These recommendations were delivered to the German Data Forum (RatSWD) in September 2008 and were published in 2009 (RatSWD 2009).

##### *4.2 Recommendations of the working group*

###### **4.2.1 Short- and medium-term recommendations**

In the view of the working group, deficiencies in the existing criminal statistics system can only be resolved effectively if the official data on crime, convictions, and prisoners are linked together in a single statistical database. Overhauling the criminal statistics system therefore means, first and foremost, creating such a database. This, however, will presumably only be accomplished in the long term. The working group therefore outlined a number of measures that could be implemented in the transition period, enabling short- and medium-term improvements of the data situation in the field of crime and crime control. These improvements serve the purpose of rectifying acute problems and are also – at least partly – necessary intermediate steps for the establishment of a statistical database. The recommendations differentiate between short- and medium-term measures that would expand and improve the existing criminal statistics system, and long-term measures that would fundamentally overhaul the system:

---

7 [http://www.ratswd.de/download/veranstaltungen/Memorandum\\_KriminalWS.pdf](http://www.ratswd.de/download/veranstaltungen/Memorandum_KriminalWS.pdf)

- (1) Measures to improve the existing components of the criminal statistics system:
  - Expanding the crime and criminal justice statistics system to include periodic crime and victimization surveys.
  - Creating a new statistic on suspects facing preliminary proceedings by public prosecutors based on data from the Central Public Prosecutors Procedural Register (ZStV, *Zentrales Staatsanwaltschaftliches Ver-fahrensregister*).
  - Creating a periodic reconviction statistic based on register data, which could enhance the ongoing research on criminal recidivism.
  - Expanding the Prison Statistics, which are currently collected on a single fixed date each year, to include entry and release statistics.
- (2) Measures to fundamentally overhaul the system:
  - Ensuring that personal data for the Criminal Justice Statistics is collected and compiled nationwide.
  - Providing pseudonymized individual-level data records from the Criminal Justice Statistics that have been checked for plausibility to the Federal Statistical Office.
  - Improving the compatibility among the Criminal Justice Statistics at the individual level as well as with the PCS.
  - Supplementing the criminal offense code with criminological-crimin-alistic characteristics including a severity index.
  - Providing greater flexibility in data collection and data preparation by supplementing the original dataset with additional information on spe-cific regions and/or time frames. These data may be collected by the authorities of the *Länder* or by a research network of statistical authori-ties and scientific researchers. The additional data can help answer existing questions or act as a trial run for future changes to the original dataset.
  - Expanding the data available from the Research Data Centers of the Federal Statistical Office and the Statistical Offices of the German *Länder* for use in all individual-level statistics on crime and criminal justice.
  - Publishing time series in digital spreadsheet form allowing for subse-quent processing.
  - Periodical publishing of Reports on Crime and Crime Control.

Carrying out at least some of these recommendations will require the passage of a federal law on individual-level statistics on criminal justice. This is the only way that will ensure that data collection and preparation is completed nationwide, that the budget is maintained, and that the pseudonymized individual-level data records of the statistics for criminal justice are delivered successfully to the Federal Statistical Office.

The Criminal Justice Statistics on legal proceedings (compiled by the public prosecutors and criminal courts) should remain unchanged. Their goals are predominantly administrative in nature, for example, to calculate the manpower needed in the judicial system.

#### 4.2.2 Long-term recommendations

Admittedly, the fundamental problems of the existing criminal statistics system will not be resolved through these short- and medium-term improvements to the data situation. The working group suggests replacing the previous system with a statistical database containing all judicial decisions of relevance to criminological research, linked with pseudonymized individual data. The problems with the current crime statistics system can be solved through this conversion. It also offers a foundation for new federal statistics collected on a regular basis, particularly regarding the preliminary proceedings of public prosecutors, remand custody, execution of sentences, and recidivism. The working group is aware, however, that this proposed system conversion will require significant organizational and technical advances. For this reason, the database presumably cannot be created until the longer term.

In an optimized criminal statistics system, the police data should also be integrated into the new database. However, due to the organizational separation of the PCS and the Criminal Justice Statistics, it is likely that this integration will not be possible in the near future. A more efficient comparison of police and court data can be achieved by harmonizing the criminal offense data collection and measurement standards.

In the opinion of the working group, the judiciary does not need to carry out any additional data collection for the criminological-statistical database system. Instead, data that have been collected for other purposes and are therefore generally available electronically should be made accessible for statistical analysis. To this end, the working group proposes that not only should the Justice Administration report to the Central Public Prosecutors Procedural Register and the Federal Central Register, but also that extensive parallel data be provided to the official statistical agencies as has been put into practice already in several *Länder*.

The working group believes that to comprehensively optimize the current criminal statistics system in Germany, the following requirements have to be met in order to enable the required system conversion:

- The existence of possibilities for identification or linkage that enable personal classification through a pseudonymized encryption process.
- Pseudonymization of the individual data records, plausibility check at *Länder* level, and provision of pseudonymized individual data to the federal level.
- Storage of the pseudonymized individual data records in a protected database at the federal level and personal linkage of the data.
- Independence of the data processing sites.

The implementation of these basic requirements calls for the establishment of a federal law:

- governing the delivery, processing, linkage, and storage of pseudonymized individual-level criminal justice data in a database,
- governing the regular compilation of the (improved) criminal justice statistics based on this dataset,
- regulating the scientific access to the individual data and the integration of the data into the institutional framework through further development of the database.

## 5. European and international developments

### 5.1 *European developments*

#### 5.1.1 *European Sourcebook of Crime and Criminal Justice Statistics*

In 1993, the Council of Europe charged a group of specialists with the preparation of a feasibility study concerning the collection of crime and criminal justice data for Europe.<sup>8</sup> The first study, which was limited to ten countries, was met with a positive response. In 1995, the expert group was therefore expanded; the first publication of the European Sourcebook of Crime and Criminal Justice Statistics in 1999 contained information from 36 European countries, covering the period 1990 to 1996.<sup>9</sup> The second edition was sponsored by the governments of Switzerland, Great Britain, and the Netherlands and published at the end of 2003.<sup>10</sup> It contained data from 40 European countries between 1995 and 2000. The third issue, which appeared in 2006, covers the period between 2000 and 2003 with data from 37

---

<sup>8</sup> <http://www.europeansourcebook.org/>

<sup>9</sup> [http://www.europeansourcebook.org/sourcebook\\_start.htm](http://www.europeansourcebook.org/sourcebook_start.htm)

<sup>10</sup> [http://www.europeansourcebook.org/esb2\\_Full.pdf](http://www.europeansourcebook.org/esb2_Full.pdf)

countries.<sup>11</sup> The publication of the fourth edition has been announced for 2009 and will cover the time period 2004 till 2007.

#### 5.1.2 The Hague Programme 2004 and its implementation

*Hague Programme 2004 and EU Action Plan 2006–2010, “Developing a Comprehensive and Coherent EU Strategy to Measure Crime and Criminal Justice.”*

The establishment of a comprehensive European crime statistics system has been recommended repeatedly. The most significant effort in this direction was the Hague Programme,<sup>12</sup> accepted by the European Council in 2004. The Council welcomed

“the initiative of the Commission to establish European instruments for collecting, analyzing and comparing information on crime and victimization and their respective trends in Member States, using national statistics and other sources of information as agreed indicators.”<sup>13</sup>

The communication from the commission “Developing a Comprehensive and Coherent EU Strategy to Measure Crime and Criminal Justice: An EU Action Plan 2006–2010,”<sup>14</sup> maintains that “one of the main deficiencies in the area of Justice, Freedom and Security is still the lack of reliable and comparable statistical information.”<sup>15</sup> The goal is to assemble “statistical information on crime (including victimization) and criminal justice at European Union level,”<sup>16</sup> which should enable the comparison between Member States and their regions. For this purpose, the “available national data will be collected and quality-assessed to form the first Community statistics on crime and criminal justice”<sup>17</sup> in the short term, and in the medium term the attempt will be made to collect data “in a harmonized manner.”<sup>18</sup> The Commission proposed to establish an expert group representing data users and a second working group should be formed to represent data producers.

---

11 [http://www.europeansourcebook.org/esb3\\_Full.pdf](http://www.europeansourcebook.org/esb3_Full.pdf)

12 Programme for the strengthening of Freedom, Security and Justice in the European Union ([http://ec.europa.eu/justice\\_home/doc\\_centre/doc/hague\\_.programme\\_en.pdf](http://ec.europa.eu/justice_home/doc_centre/doc/hague_.programme_en.pdf)).

13 Hague Programme (note 12): 25.

14 <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2006:0437:FIN:EN:PDF>

15 EU Action Plan 2006-2010 (note 14): 3.

16 EU Action Plan 2006-2010 (note 14): 2.

17 EU Action Plan 2006-2010 (note 14): 3.

18 EU Action Plan 2006-2010 (note 14): 4.

*Actions for the implementation of the Hague Programme and the EU Action Plan 2006–2010*

*Expert group of the Directorate-General for Justice, Freedom and Security within the EU Commission (DG JLS):* pursuant to the EU Action Plan 2006–2010, the DG JLS set up a group of experts on the policy needs for data on crime and criminal justice (PNDCCJ).<sup>19</sup> Among other things, this group had the task of creating an indicator list on the comparison of criminal statistics in Europe and of developing the content requirements for a Europe-wide population survey on crime rates and safety concerns. The expert group identified human trafficking and money laundering as issues on which urgent comparative data should be gathered across Europe.<sup>20</sup>

*Working group on Statistics on Crime and Criminal Justice:* also in 2006, the Directors of Social Statistics of the National Statistics Departments in Europe (DSS) created a representative working group for data producers “Statistics on Crime and Criminal Justice.”<sup>21</sup> It was tasked with the responsibilities laid out in the Hague Programme 2004 of supporting in data collection, analysis, and comparison in the field of crime and criminal justice in the Member States, and ensuring that the EU Action Plan is implemented effectively by working closely with other supranational organizations that are already active in the field of crime and judicature, for example the European Sourcebook Group,<sup>22</sup> the United Nations Office on Drugs and Crime (UNODC),<sup>23</sup> and the United Nations Economic Commission for Europe (UNECE).<sup>24</sup> Additionally, the working group will pass their work and results onto the statistical systems of the Member States. Moreover they will be required to appoint specific task forces<sup>25</sup> and to cooperate closely with the expert group of the Directorate-General for Justice, Freedom and Security within the EU Commission (DG JLS) (see above).

*Statistics in focus: Crime and Criminal Justice:* it was proposed in the Hague Programme that Eurostat should use European instruments for the collection, analysis, and comparison of information on crime and victimization and also draw on the developments of the individual Member

---

19 Commission Decision of 7 August 2006 setting up a group of experts on the policy needs for data on crime and criminal justice, Official Journal L 234, 29/08/2006 P. 0029 – 0032. <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2006:234:0029:01:EN:HTML>

20 <http://www.crime-prevention-intl.org/filebin/Generating%20Links%20for%20Website/Other%20ICPC%20Events/Rencontre%20des%20Observatoires%20de%20Paris/Liens/Olivier.Bardin.Paris.ENG.pdf>

21 [http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2006:234:0029:01:EN:HTML: 2](http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2006:234:0029:01:EN:HTML:2)

22 <http://www.europeansourcebook.org/>

23 <http://www.unodc.org/unodc/en/about-unodc/index.html>

24 <http://www.unece.org/Welcome.html>

25 Currently two task forces exist on “victimization” and “crime data availability.”

States. Additionally, they were to use national statistics and other information sources as agreed indicators. In 2007, the first copy of *Statistics in Focus: Crime and Criminal Justice*<sup>26</sup> was published; in the meantime the first updated version has been released.<sup>27</sup>

The publication contains the most important results of the Member States' "Crime Data Request" led by Eurostat. While the publication only identifies trends, absolute numbers are available on the Eurostat website.<sup>28</sup>

*Europe-wide Population Survey on Crime Rates and Safety:* in the implementation of the EU Hague Programme and the EU Action Plan 2006–2010, Eurostat plans to conduct a European victimization survey in 2010/2011 on victimization and general safety issues. In preparation, a survey of previous victim surveys conducted in Europe was created. On this basis – and using the manual on victim interviews developed by the UN – a questionnaire were developed and tested.

The European Institute for Crime Prevention and Control affiliated with the United Nations (HEUNI)<sup>29</sup> was commissioned to conceptualize the design and the questionnaire material (Heiskanen and Viuhko 2007; Thomas 2007). The individual methods of such a population survey should be tested using different surveying methods in the individual Member States. Germany wants to participate in the pilot survey; therefore the Federal Statistical Office, four of the Statistical Offices of the German *Länder*, and the BKA collectively developed an operating concept that they submitted to Eurostat.

## 5.2 International developments

*International Crime Victimization Survey (ICVS) and International Self-Report Delinquency Study:* the major advantage of crime surveys is that they have a standardized inventory and are independent of national crime definitions, convictions, and registration practices. They are therefore suitable for country comparisons.

In the 1980s the International Crime Survey (ICS) was established, which has been known as the ICVS since its renaming in 1996. The survey has been repeated using the same instruments and most of the same methods<sup>30</sup> four times since 1989 – that year also with German participation.<sup>31</sup>

26 [http://epp.eurostat.ec.europa.eu/cache/ITY\\_OFFPUB/KS-SF-07-015/EN/KS-SF-07-015-EN.PDF](http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-SF-07-015/EN/KS-SF-07-015-EN.PDF)

27 [http://epp.eurostat.ec.europa.eu/cache/ITY\\_OFFPUB/KS-SF-08-019/EN/KS-SF-08-019-EN.PDF](http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-SF-08-019/EN/KS-SF-08-019-EN.PDF)

28 [http://epp.eurostat.ec.europa.eu/portal/page?\\_pageid=1996,45323734&\\_dad=portal&\\_schema=PORTAL&screen=welcomeref&open=/&product=EU\\_MASTER\\_crime&depth=2](http://epp.eurostat.ec.europa.eu/portal/page?_pageid=1996,45323734&_dad=portal&_schema=PORTAL&screen=welcomeref&open=/&product=EU_MASTER_crime&depth=2)

29 <http://www.heuni.fi/index.htm>

30 <http://rechten.uvt.nl/icvs/>

Germany participated again in the fifth survey that was carried out in 2004/05 (Van Dijk et al. 2008). The sixth survey is planned for 2009.

An “International Self-Report Delinquency Study” took place in 1982 (Junger-Tas et al. 1994). The second International Self-Report Delinquency Study occurred with German participation.<sup>32</sup>

*Comparative international crime and criminal justice statistics:* The International Criminal Police Organization (ICPO, *Interpol*)<sup>33</sup> gathers data from the police crime statistics of their Member States and then summarizes these data at regular intervals. These reports were previously available upon request and for a few years they were even universally accessible on the Internet. However, for the past few years, they have only been available for (internal) official use. In 2006, the 75th General Assembly passed a resolution<sup>34</sup> that brought an end to the International Criminal Police Statistics.<sup>35</sup>

*United Nations Survey on Crime Trends and the Operations of Criminal Justice Systems (World Crime Survey):* following international trends, the United Nations compiles data on international crime developments as well as sentencing and conviction by law enforcement agencies.<sup>36</sup> The United Nations Survey on Crime Trends and the Operations of Criminal Justice Systems (World Crime Survey) has been executed in eleven phases since 1970. The accuracy and reliability of this survey is naturally dependent on how accurately the comprehensive questionnaire is filled out.<sup>37</sup>

---

31 At the first ICVS in 1989, 14 countries including Germany participated. The following surveys took place in 1992 with the participation of 33 countries, and in 1996 and 2000, 48 countries participated. Germany did not participate in these latter two surveys. Germany only took part again in the fifth survey of 2004/2005.

32 <http://www2.jura.uni-hamburg.de/instkrim/kriminologie/Projekte/ISRD2/ISRD2.html>. The final report has been announced for the summer of 2008.

33 <http://www.interpol.int/>

34 Resolution No AG-2006-RES-19.

35 <http://www.interpol.int/public/ICPO/GeneralAssembly/AGN75/resolutions/AGN75RES19.asp>

36 <http://www.unodc.org/unodc/en/data-and-analysis/United-Nations-Surveys-on-Crime-Trends-and-the-Operations-of-Criminal-Justice-Systems.html>

37 For the comparison of data quality and individual internationally comparative analyses, see the second Periodical Report on Crime and Crime Control (Federal Ministry of the Interior and Federal Ministry of Justice 2006: 39ff.).



## References:

- Albrecht, H.-J. (2001): Informationsfeld "Rechtspflege/Gerichtsbarkeit/innere/äußere Sicherheit". In: Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik (KVI) (Ed.): Wege zu einer besseren informationellen Infrastruktur. Baden-Baden.
- Federal Ministry of the Interior and Federal Ministry of Justice (Ed.) (2001a): First Periodical Report on Crime and Crime Control. Abridged Version. Berlin. [http://www.uni-konstanz.de/rtf/ki/Download\\_Abridged\\_Version.pdf](http://www.uni-konstanz.de/rtf/ki/Download_Abridged_Version.pdf). [Last visited: 02/19/2010].
- Federal Ministry of the Interior and Federal Ministry of Justice (Ed.) (2001b): First Periodical Report on Crime and Crime Control. Berlin.
- Federal Ministry of the Interior and Federal Ministry of Justice (Ed.) (2006a): Zweiter Periodischer Sicherheitsbericht. Berlin. <http://www.uni-konstanz.de/rtf/ki/links.htm#PSB2>. [Last visited: 02/19/2010].
- Federal Ministry of the Interior and Federal Ministry of Justice (Ed.) (2006b): Second Periodical Report on Crime and Crime Control in Germany. Abridged Version. Berlin.
- Heinz, W. (2002): Final Report by the Working Group of the Ministry of the Interior and the Ministry of Justice on the Regular Execution of Victim Surveys. [Unpublished paper].
- Heinz, W. (2003): Soziale und kulturelle Grundlagen der Kriminologie – Der Beitrag der Kriminalstatistik. In: Dittmann, V. and Jehle, J.-M. (Eds.): Kriminologie zwischen Grundlagenwissenschaft und Praxis. Mönchengladbach.
- Heinz, W. (2008): Kriminalstatistik – quo vadis? In: Festschrift für K. Tiedemann. Cologne.
- Heiskanen M. and Viuhko, M. (2007): European Victimization Survey. In: Aromaa, K. and Heiskanen, M. (Eds.): Victimization Surveys in Comparative Perspective. Papers from the Stockholm Criminology Symposium. <http://www.heuni.fi/uploads/ojyw1estdwh5b.pdf>. [Last visited: 02/19/2010].
- Junger-Tas, J./Terlouw, G.-J. and Klein, M.W. (Eds.) (1994): Delinquent behavior among young people in the western world: first results of the international self-report delinquency study. Amsterdam.
- Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik (KVI) (Ed.) (2001): Wege zu einer besseren informationellen Infrastruktur. Baden-Baden.
- Rat für Sozial- und Wirtschaftsdaten (RatSWD) (Ed.) (2009): Optimierung des bestehenden kriminalstatistischen Systems in Deutschland. Baden-Baden.
- Thomas, G. (2007): Development of a EU Victimization Module. In: Aromaa, K. and Heiskanen, M. (Eds.): Victimization Surveys in Comparative Perspective. Papers from the Stockholm Criminology Symposium. <http://www.heuni.fi/uploads/ojyw1estdwh5b.pdf>. [Last visited: 02/19/2010].
- Van Dijk, J.J.M./van Kesteren, J.N. and Smit, P. (2008): Criminal Victimization in International Perspective. Code findings from the 2004-2005 ICVS and EU ICS. The Hague. [http://rechten.uvt.nl/icvs/pdf/ICVS2004\\_05.pdf](http://rechten.uvt.nl/icvs/pdf/ICVS2004_05.pdf). [Last visited: 02/19/2010].



## 6.6 Environment

*Cornelia Ohl and Bernd Hansjürgens*

Contact:

Cornelia Ohl  
Bernd Hansjürgens  
Helmholtz Centre for Environmental Research – UFZ  
Department of Economics  
Permoserstraße 15  
04301 Leipzig  
Germany  
e-mail: [cornelia.ohl\[at\]ufz.de](mailto:cornelia.ohl[at]ufz.de)  
[cornelia.ohl\[at\]uba.de](mailto:cornelia.ohl[at]uba.de)  
[bernd.hansjuergens\[at\]ufz.de](mailto:bernd.hansjuergens[at]ufz.de)

## **Abstract**

Many environmental problems are large in scale, both geographically and temporally. A natural implication of this is that numerous environmental events and consequences often tend to coincide with one another, therefore making inferences about causal relationships between humans and nature is difficult at best (“transparency challenge”). Consequently, we see a need for innovative analytical methods and modeling approaches to supplement the traditional monitoring-based approach to environmental policy. These methods and modeling approaches should make it possible to capture different degrees of uncertainty, which are generally beyond the scope of the control variables used in traditional monitoring activities. Moreover, due to the difficulty associated with distinguishing the boundaries between natural and social systems, the monitoring approaches should be based on the collection and connection of data from different fields. For this, comparable and often very large datasets are needed, (“availability challenge” and “compatibility challenge”). Even if these obstacles are overcome, data processing will remain a very complex and time-consuming task. It is therefore crucial that the data infrastructure is user-friendly. We see an advantage here in using the Geographical Information Systems (GIS) technology as well as a nested structure for data provision. This structure should aid the upwards and downwards scaling of information as well as facilitate access to data by the relevant parties – polluters, victims, and regulators (“connectivity challenge”).

Keywords: coincidence of causes and impacts, transparency, availability, compatibility, connectivity, GIS technology, nested structure for data provision

## **1. Research questions**

At present, various data sources indicate that human interferences in nature have reached unprecedented levels (Vitousek et al. 1997; Nelson et al. 2006). Disturbances in ecological systems have increased in magnitude, and impact not only ecosystems, but also human wellbeing (IPCC 2007; MEA 2005). Unsurprisingly, this has led to increased questioning of contemporary management strategies. In order to effectively monitor human impact on his surroundings, new monitoring approaches and strategies are required. Ideally, these strategies should better capture the interplay between humans and nature in terms of the associated varying degrees of uncertainty and unpredictability involved.

The fact that many environmental problems today are large in scale both geographically and temporally poses several problems for the design of data provision. Specifically, these problems pertain to the collection, processing, and accessibility of information. Reasons for this are: (1) the fact that a majority of environmental problems are long-term in nature, which frequent-

ly makes it impossible to distinguish between overlapping causes and impacts. (2) Furthermore, many environmental problems are characterized by “true uncertainty.” This means that neither the expected damages and costs nor the probability of their occurrence are adequately understood. In turn, (3) this makes it difficult to monitor policy interventions due to changes in relevant state variables and (4) undermines the evaluation of policy success or failure. Hence, any monitoring activity aiming to capture changes in complex environmental phenomena is doomed to fail. Thus, the critical questions to answer are: how does one monitor changes in complex environmental phenomena characterized by human interference and human-nature feedbacks? How does one measure the success of policy interventions given that environmental problems are large-scale and long-term?

To answer these questions, we see a need for innovative analytical methods and modeling approaches that supplement the traditional monitoring-based approach. Traditional monitoring is based on indicator-orientated environmental assessments of typically a single (often well-known) source or pollutant that causes specific ecosystem disturbances. Analytical methods and modeling approaches are required in order to better understand correlations between multiple interferences in natural and societal systems as well as the interplay of different sources and pollutants. We call this the “transparency challenge.” It is the challenge of separating multiple interferences at different levels of interplay (e.g., the level of drivers and responses). As a prerequisite for defining policy responses aimed at halting undesired environmental changes, these innovative methods and approaches require descriptive informational inputs about the natural system being assessed. This includes information not only about internal pressures affecting the state of the ecosystem, but also about external pressures, such as the effects of various sectors of the economy, local planning bodies, etc. We label this the “availability challenge.” It is the challenge of delivering sufficient information about different facets of a problem to researchers, policy-makers, and the public. Here, problems may arise if data are non-existent or not accessible. In this regard, it is important to note that environmental problems are of a particular nature; they are characterized by problems of fit, interplay, and scale (Young 2002). The notion of fit refers to the natural and socio-economic boundaries of an environmental problem. While natural boundaries are determined by the “natural properties” of an ecosystem, for example, the boundaries of a river basin, the socio-economic boundaries are mainly administrative ones, for example, national, regional, or local governmental units. From the perspective of providing adequate data, this leads to a misfit such that the environmental problem under concern cannot be adequately captured in a quantitative dimension. We call this the “compatibility challenge.” It is the challenge of avoiding a mismatch of available datasets. Problems of interplay refer to the fact that many environ-

mental problems are cross-cutting by nature. This requires that data from several policy fields be combined. Making proposals for alternative land use, for instance, requires information on agriculture, urbanization, the water sector, and other environmental media. In most cases, existing databases have not been developed for cross-cutting research questions like this, so major changes with respect to data processing have to be undertaken. Finally, problems of scale refer to the need to scale data up and down. Many environmental data are gathered on a scale different from that where policy recommendations are developed. This leads us to formulate a fourth challenge, the “connectivity challenge.” This is the challenge of combining the available datasets in such a way that information is accessible from different perspectives (e.g., from the victim’s, policy-maker’s, and pollutant’s points of view). These challenges require new and innovative methods of data management.

## **2. Status quo: Databases and access**

In recent years, considerable progress has been achieved in the provision of both natural and social science data on environmental issues. Today, it is undisputed that social and behavioral science data are complementary and supplement natural science data. At an international level, four major data approaches can be identified (Ohl et al. 2009):

- (1) *The media approach*: focused on environmental components, such as air, land, water, and the human-made environment;
- (2) *The stress-response approach*: focused on human impacts on the environment and subsequent transformation (“responses”) of environmental systems;
- (3) *The resource accounting approach*: focused on the natural resources flow beginning at extraction, then following different resource uses throughout the lifetime of a product, and ending at the final return of the resources (e.g., as emissions, wastewater) into the environment;
- (4) *The ecological approach*: based on models, monitoring techniques, and ecological indices. This approach draws on the notion of pressures, states, and responses (PSR) with regard to data organization, but applies these concepts only to ecological zones within a country. For example, the Geographical Information Systems (GIS) use the ecological approach.

Additionally, different combinations of these approaches are used on all scales of environmental statistics (local, regional, national) (Ohl et al. 2009):

- FDES: a Framework for the Development of Environmental Statistics, developed by the United Nations Statistical Office;
- PSR: Pressure-State-Response framework, developed by the Organization for Economic Co-operation and Development (OECD);
- DSR: Driving forces-State-Response framework, developed by the Commission of Sustainable Development;
- DPSIR: Driving forces-Pressure-State-Impact-Response framework, used by the European Environment Agency (EEA) and the Statistical Office of the European Communities (Eurostat).

The use of these approaches and frameworks has led to extensive data on all kinds of environmental issues. These occurred not only in Europe, but all over the world. To some extent, these approaches complement each other, as each emphasizes different aspects of an environmental issue. The differences in viewpoint, however, are sometimes confusing. For example, the collection of data on drivers and pressures in abstract of one another is only used in the DPSIR framework. The other frameworks do not differentiate between them.

Despite this disparity among approaches, the overall experience in collecting and reporting environmental data has led to the development of several useful environmental indicators since the 1970s. These indicators allow reporting on, for example, environmental conditions, environmental performance, and progress towards sustainable development. These indicators are judged as powerful cost-effective tools for tracking environmental progress, providing policy feedback, and measuring environmental performance (OECD 2003). Their development has catalyzed fruitful cooperation among a great number of countries and international organizations, for example, between the OECD and the United Nations Statistics Division (UNSD), the UN Commission for Sustainable Development (UNCSD), the United Nations Environment Programme (UNEP), the Commission of the European Communities, Eurostat, and the European Environment Agency (EEA).

In addition, considerable progress has been made regarding the development of a System of Integrated Environmental Economic Accounting (SEEA). This has occurred at both the UN level and at respective national levels, for example, in the German Environmental Accounting Framework (UGR, *Umweltökonomische Gesamtrechnung*). The UGR was developed in the 1990s and delivers the most comprehensive framework for capturing the relationships between environment and economy today. Both approaches, the SEEA and the UGR, are characterized by an integrative perspective that makes use



of common concepts, definitions, and classifications in order to allow for the direct observation of links between economic and environmental development. Importantly, these approaches serve as a basis for providing indicator-based information to policy-makers and the public. Moreover, the integrated accounting approaches allow for drawing conclusions about the macro-economic costs of policy measures by modeling sector-specific economic and environmental behavior under certain policy constraints. These approaches are currently being evaluated and revised by the UN Committee of Experts on Environmental Economic Accounting and Statistics (UNCEEAA), with the aim of developing an international statistical standard.

Against this background, deficits in data provision are hardly found on the macroeconomic level. What is missing is the provision of adequate datasets on the microeconomic level. If made available, these micro-level datasets could be linked with already readily available macro-level data to enhance our understanding of the vulnerability of individuals and social groups subject to environmental change on the level of small scale regional units.

### **3. Future developments**

#### *3.1 Data provision*

In addition to the information provided by statistical institutions and other organizations, there are several other networks responsible for data provision. One of these is, for example, the Global Earth Observation System of Systems (GEOSS), which, within the next decade, is expected to provide a further large amount of new datasets. These include several products such as maps on river systems, infrastructure, land cover, and land use, all of which are expected to be available for common use. To interpret and use these products for society's benefit, the earth observation data need to be linked to social science information on human related drivers and consequences of change. Currently, there are two problems associated with data provision in GEOSS. First, socio-economic data providing this kind of information is very often on administrative scales that differ from natural scales, creating a problem of fit (see above). Second, socio-economic data and indicators are rarely delivered and visualized in maps. Nonetheless, it should still be noted that some progress has been made in the technical support of this kind of data provision, especially since GIS<sup>1</sup> technology has improved the effectiveness and analytic power of traditional mapping.

---

1 GIS application tools support users in analyzing spatial information (i.e., data that refers to or is linked to a specific location), in editing data, and in visualising the results of operation

Today, in several fields of application, GIS not only provides maps on socio-economic developments in space and time, but also supports analyses of social science data for decision-making. For marketing purposes, for instance, demographic information is used to determine how many individuals with a certain socio-economic classification (e.g., age, sex, or income) live in a given spatial area. Another prime example is the CompStat approach used in New York City, which uses GIS for crime mapping and analysis (e.g., crime forecasting and geographic profiling) to formulate strategies, target resources, and evaluate crime reduction programs. Data held by GIS may also be used as a spatial decision-support system. In the US, time-specific population data, which deliver insight into humans' daily routines, are used to track and model patterns of commuter behavior. Projecting these data forwards into the future is helpful in assisting the local planning bodies in analyzing and testing different types of policy decisions.

In the field of the environment, the most prominent example is the use of GIS to understand the impacts of global climate change. So far, however, the focus has mainly been on combining various maps and satellite information sources to simulate interactions of complex natural system phenomena (e.g., the impacts of climate change on coastal areas, including flooding due to sea-level rise and storm erosion). According to these data, the exposure of individuals, societal groups, or regions to climate change risks and impacts can be visualized. A future challenge will be including anthropogenic factors in order to better understand the capacity of the entities considered to cope with climate change impacts. Relevant questions in this regard are: which individuals or social groups are affected by global climate change? What is their regional distribution (e.g., within the boundaries of an urban agglomeration)? What are the housing conditions? Are individuals able to protect their homes against flooding or to cope with flooding events? Is it possible to combine global climate change data with data on social segregation? Can changes in lifestyle or socio-economic adaptation measures be captured? The final goal of adding these data to the existing global climate change data is to get a deeper understanding of the vulnerability of individuals, social groups, societies, and regional units. This includes data on both the exposure of "elements at risks" as well as coping capacities.

Further questions arise with respect to the measurement of the success of policy responses: how should a new type of regulation consider variables on the state of the natural and social system? Who will be affected by current and future regulations and how? Answering these questions should deliver the blueprint for building modern data provision infrastructure. And, of course, this infrastructure will have to be updated whenever a new policy problem materializes. Here, the challenge is that for the observation of newly

---

in maps. GIS can, for example, be used for urban planning, resource management, and environmental impact assessment.

emerging environmental problems, the existing infrastructure needs to be flexible enough to adapt to and be merged with newly emerging demands for data provision. A second important challenge is to identify overlaps of natural and social systems, especially with regard to the social entities affected. This concern is related to the problem of interplay (see above) as well as correlations between the new and past chain of causes and impacts.

### *3.2 Data usage*

The most important deficit in the field of data usage is the improper provision of information for the implementation of policy responses. The provision of data does not sufficiently take into account the needs of data users. This primarily holds for trans-boundary and global environmental phenomena (Neßhöver et al. 2007: 120), but also at the regional and local level. To overcome this shortcoming, the design of monitoring activities needs to stringently take policy considerations and users' needs into account. Questions concerning design need to be asked in a targeted manner: which information is required for which purpose, at which point in time, and by which user (e.g., at which governmental level)? Very often, data collection, processing, and publication are driven by the providers, the "supply side." It will be, however, crucial to strengthen the interests of users in the process of collecting and processing data and thereby strengthen the "demand side."

The environmental data available is often insufficient not only for policy evaluation, but also for public communication purposes. One important goal of the collection and distribution of environmental data is to inform the general public. In order to achieve this goal, information has to be prepared in a way that stakeholders who are not experts in a particular environmental field are able to understand and interpret the data.

However, public participation and the involvement of user groups can even go a step further. To foster public involvement in policymaking as well as to promote the goals of nongovernmental organizations, grassroots groups, and community-based organizations, data infrastructure should broaden its view to public participation. In this regard Public Participation GIS (PPGIS, Sieber 2006) can be used as a supportive tool. Ghose (2001) reports a case study where residents of an inner city neighborhood became active participants in building a community information system. The participants learned to access public information and create and analyze new databases derived from their own surveys. In this manner, participants became engaged in city management and the formation of public policy. Use of PPGIS is motivated by the expectation that access to information is the key to more effective government and community empowerment. As a top-down approach, PPGIS could also be used to analyze the spatial differences in access to environmental services (e.g., with reference to the social and economic background

of relevant actors) and thus support making adjustments and improvements in environmental management.

### 3.3 Data access

The vast amount of data provided by institutions and organizations is easily accessible via the Internet. However, datasets are often dispersed and disconnected and thus inconvenient for users to handle. In cases where datasets are centrally archived, for example, on the homepages of the United Nations Framework Convention on Climate Change (UNFCCC) or the Convention on Biological Diversity (CBD), the sheer amount of information available often exceeds the time constraints of the users seeking it.<sup>2</sup> The speed of progress in computer technology and widespread Internet access, together with the complexity of the problems under consideration – especially if they are global in scope – are some of the reasons why desired information is not always easily accessible. This holds true not only for third-party users in the general public, but also for individuals responsible for the provision and analysis of the datasets. Hence, it is unsurprising that the relevant data suffer from time lags in provision and do not qualify as up to date.

Thus, although data provision has considerably improved in recent years due to technological progress in the information sector, the main factor hampering information processing is human: too much complex information for a normal human being to process, and time constraints create bottlenecks. To deploy and process more of the information provided by administrative sources, it is thus necessary to assist the users with improved search functions and an infrastructure that allows for individual ways of data connection. One promising route to follow in this regard is again the development of a GIS based system of data storage and processing.

To fully utilize the societal benefits of environmental data, it is crucial to share the data across national and international administrative boundaries. In this regard, the GEOSS data-sharing principles could work as a model for future developments in national and international data-sharing. In recognition of relevant international instruments and national policies and legislation, GEOSS will support the full and open exchange of data, metadata, and products, not only within the GEOSS community, but also beyond. For research and education, all shared data, metadata, and products will be provided free of charge or at no more than the cost of production. For other users, these

---

2 One example is reporting by CBD signatories on measures undertaken and their effectiveness in accordance with article 26 of the Convention. So far there are 191 CBD parties, of which 143 have delivered the third national report (NR3); see <http://www.cbd.int/reports/> (accessed November 30, 2008). Going through all these reports to find country-specific information on a particular measure is an extremely time-consuming task.

data will be provided at minimum cost. Use of the data or products need not necessarily imply agreement with or endorsement of the purpose behind the gathering of the data, which will be made available with minimum time delay.

At the local and individual level, the security of environmental data may remain a problem for the social sciences. While data collection on a very small scale is usually not a problem in the natural sciences, the collection of such data in the socio-economic fields can become a problem if individuals, households, or companies can be identified due to the small number of units in the sample. Here, the legal protection of the private sphere of individuals, households, and companies may lead to conflicts with research interests.

#### **4. Future developments: European and international challenges**

Despite important progress in the field of international environmental statistics, differences among countries remain. In order to make progress in providing environmental data tailored towards policy needs, we must establish closer links between the data gathered on natural systems and the data on social systems at different scales. In this respect, linking national accounts with international datasets seems to be most important. A nested structure of data provision that provides the datasets from various points of view seems appropriate. These include:

- Polluters: focusing on, e.g., consumption behavior and production processes.
- Victims: focusing on, e.g., the consumption of harmful goods, or vulnerability of specific sectors in the economy due to climate change.
- Regulators: focusing on an inventory of policies affecting, e.g., environmental pollution behavior and reducing social vulnerabilities.

Coordinated data management by national and supra-national bodies should center on *environmentally relevant core activities*. Determination of these core activities requires an approach that includes the interests of (national) users. Such an approach would facilitate agreement on the objectives of data gathering and sharing as a prerequisite for developing a common data infrastructure. Guiding questions in this regard are:

- What are the most important environmental problems that need to be solved on a supranational level (climate change, biodiversity loss, water scarcity, deposition of nuclear waste, etc.)?
- Which state variables describe the problem under consideration (e.g., emission levels, damage costs, stock of resources)?
- What are the key variables that require monitoring and policy control (e.g., sectors, inputs, outputs)?
- What are the most important channels for transferring impacts from one administrative unit (governance level) to another (e.g., import and export of goods, unidirectional or reciprocal externalities, etc.)?
- Within which time horizon do the problems need to be solved and should a policy phase-out take place (considering delays as well as persistence and irreversibility of causes and impacts)?
- Which policy measures already affect or are expected to affect the problem under consideration?

With regard to the organizational infrastructure, an improved systematic horizontal and vertical integration of datasets from different types of administrative, research, and business units is urgently needed. The key aim of horizontal integration is to develop standards for the integration of important private (business) and project-related research data in official accounts at all administrative levels. The key aim of vertical integration is to derive national accounts data from datasets collected on the lower (sub-national) administrative units and vice versa. This requires developing ways of combining electronic surveys with new sampling techniques and/or algorithms that are capable of exploiting data at different levels of generalization. This in turn involves cross-linking statistical data, including its combination with text- and image-based information available from different sources. For this purpose, it will be crucial to develop a sophisticated infrastructure for data storage and provision (e.g., development of statistical and machine learning algorithms that have the capacity to cope with massive amounts of data, development of methodologies and semantics for statistics, integrated with metadata construction and retrieval systems to handle statistical requests and improve the access to datasets). Only the future will show whether the improvements needed here will be made.

## References:

- Global Earth Observation System of Systems (GEOSS): 10-Year Implementation Plan (As adopted 16 February 2005), accessed <http://www.earthobservations.org/documents/10-Year%20Implementation%20Plan.pdf> [last visited: 02/17/2010]; see also <http://www.earthobservations.org/> [last visited: 02/17/2010].
- Ghose, R. (2001): Use of Information Technology for Community Empowerment: Transforming Geographic Information Systems into Community Information Systems. *Transactions in GIS* 5 (2), 141-163.
- Intergovernmental Panel on Climate Change (IPCC) (2007): Climate Change 2007 – Synthesis Report. IPCC Fourth Assessment Report. Cambridge: [http://www.ipcc.ch/pdf/assessment-report/ar4/syr/ar4\\_syr.pdf](http://www.ipcc.ch/pdf/assessment-report/ar4/syr/ar4_syr.pdf). [Last visited: 02/17/2010].
- Millennium Ecosystem Assessment (MEA) (2005): Ecosystems and human well-being. Washington.
- Nelson, G.C./Bennett, E./Berhe, A.A./Cassman, K./DeFries, R./Dietz, T./Dobermann, A./Dobson, A./Janetos, A./Levy, A./Marco, D./Nakicenovic, N./O'Neill, B./Norgaard, R./Petschel-Held, G./Ojima, D./Pingali, P./Watson, R. and Zurek, M. (2006): Anthropogenic Drivers of Ecosystem Change: An Overview. *Ecology and Society* 11 (2), 29. <http://www.ecologyandsociety.org/vol11/iss2/art29/>. [Last visited: 02/17/2010].
- Neßhöver, C./Berghöfer, A. and Beck, S. (2007): Weltranglisten als Beratungsinstrumente der Umweltpolitik – Eine Einschätzung des Environmental Performance Index. Marburg.
- OECD (2003): OECD Environmental Indicators: Development, Measurement and Use. Paris. <http://unpan1.un.org/intradoc/groups/public/documents/apcity/unpan015281.pdf>. [Last visited: 02/14/2010].
- Ohl, C./Bezák, P./Palarie, T.A./Gelan, A. and Krauze, K. (2009): DAPSET – Concept for Characterising Socio-Economic Drivers of and Pressures on Biodiversity. UFZ-Discussion Paper 2/2009.
- Sieber, R. (2006): Public Participation and Geographic Information Systems: A Literature Review and Framework. *Annals of the American Association of Geographers* 96 (3), 491-507.
- Vitousek, P.M./Mooney, H.A./Lubchenco, J. and Melillo, J.M. (1997): Human Domination of Earth's Ecosystems. *Science* 277 [25 July 1997], 494-499.
- Young, O. (2002): The Institutional Dimension of Environmental Change. Fit, Interplay, and Scale. Cambridge.





## TERMINOLOGY

English Name <sup>1</sup>	Original German Name
Research Data Center of the Federal Statistical Office	Forschungsdatenzentrum des Statistischen Bundesamtes
Research Data Center of the Statistical Offices of the German <i>Länder</i>	Forschungsdatenzentrum der Statistischen Landesämter
Research Data Center of the Federal Employment Agency (BA) at the Institute for Employment Research (IAB)	Forschungsdatenzentrum der Bundesagentur für Arbeit (BA) im Institut für Arbeitsmarkt- und Berufsforschung (IAB)
Research Data Center of the German Pension Insurance (RV)	Forschungsdatenzentrum der Rentenversicherung (RV)
Research Data Center of the Federal Institute for Vocational Education and Training (BIBB)	Forschungsdatenzentrum im Bundesinstitut für Berufsbildung (BIBB)
Research Data Center at the Institute for Educational Progress (IQB)	Forschungsdatenzentrum am Institut zur Qualitätsentwicklung im Bildungswesen (IQB)
Research Data Center of the German Socio-Economic Panel (SOEP)	Forschungsdatenzentrum des Sozio-oekonomischen Panels (SOEP)
Research Data Center ALLBUS at the Leibniz Institute for the Social Sciences (GESIS)	Forschungsdatenzentrum ALLBUS bei GESIS
Research Data Center "International Survey Programs" at the Leibniz Institute for the Social Sciences (GESIS)	Forschungsdatenzentrum "Internationale Umfrageprogramme" bei GESIS
Research Data Center "Voting Behavior database" at the Leibniz Institute for the Social Sciences (GESIS)	Forschungsdatenzentrum "Daten der Wahlforschung" bei GESIS
Research Data Center of the Survey of Health, Ageing and Retirement in Europe (SHARE)	Forschungsdatenzentrum des Survey of Health, Ageing and Retirement in Europe (SHARE)
Research Data Center of the German Ageing Survey (DEAS)	Forschungsdatenzentrum Deutscher Alterssurvey (DEAS)
German Microdata Lab (GML) Service Center for Microdata at the Leibniz Institute for the Social Sciences (GESIS)	German Microdata Lab (GML) Servicezentrum für Mikrodaten des Leibniz-Instituts für Sozialwissenschaften (GESIS) / MISSY
International Data Service Center at the Institute for the Study of Labor (IZA)	Internationales Datenservicezentrum des Forschungsinstituts zur Zukunft der Arbeit (IZA)

<sup>1</sup> For the purpose of standardization, the German names of the Research Data Centers have been translated into English using the American spelling.

## Frequently used Abbreviations:

### A

AES	(European) Adult Education Survey – <i>Berichtssystem Weiterbildung</i>
AFiD	Official Firm Data for Germany – <i>Amtliche Firmendaten für Deutschland</i>
ALL	Adult Literacy and Lifeskills Survey
ALLBUS	German General Social Survey – <i>Allgemeine Bevölkerungsumfrage der Sozialwissenschaften</i>
ASID	Old-age Pension Schemes in Germany – <i>Alterssicherung in Deutschland</i>
AVID	Retirement Pension Provision Schemes in Germany – <i>Altersvorsorge in Deutschland</i>

### B

BA	Federal Employment Agency – <i>Bundesagentur für Arbeit</i>
BAMF	Federal Office for Migration and Refugees – <i>Bundesamt für Migration und Flüchtlinge</i>
BAP	Employment Panel of the BA – <i>BA-Beschäftigtenpanel</i>
BBR	The Federal Office for Building and Regional Planning – <i>Bundesamt für Bauwesen und Raumordnung</i>
BCS	British Cohort Study
BHP	Establishment History Panel – <i>Betriebs-Historik-Panel</i>
BHPS	British Household Panel Study
BiB	Federal Institute for Population Research – <i>Bundesinstitut für Bevölkerungsforschung</i>
BIBB	Federal Institute for Vocational Education and Training – <i>Bundesinstitut für Berufsbildung</i>
BIJU	Learning Processes, Educational Careers, and Psychosocial Development in Adolescence and Young Adulthood – <i>Bildungsverläufe und psychosoziale Entwicklung im Jugend- und jungen Erwachsenenalter</i>
BiKS	Educational Processes, Competence Development, and Selection Decisions in Pre- and Primary School Age – <i>Bildungsprozesse, Kompetenzentwicklungen und Selektionsentscheidungen im Vor- und Grundschulalter</i>

BMBF	Federal Ministry of Education and Research – <i>Bundesministerium für Bildung und Forschung</i>
BStatG	Law on Statistics for Federal Purposes – <i>Gesetz über die Statistik für Bundeszwecke (Bundesstatistikgesetz)</i>
<b>C</b>	
CAPI	computer-assisted personal interview
CASI	computer-assisted self interview
CATI	computer-assisted telephone interview
CESSDA	Council of European Social Science Data Archives
CNEF	Cross-National Equivalent File
CSES	Comparative Study of Electoral Systems
<b>D</b>	
DDI	Data Documentation Initiative
DEAS	German Ageing Survey – <i>Deutscher Alterssurvey</i>
DESI	Assessment of Student Achievements in German and English as a Foreign Language – <i>Deutsch-Englisch-Schülerleistungen-International</i>
DFG	German Research Foundation – <i>Deutsche Forschungsgemeinschaft</i>
DGS	German Sociological Association – <i>Deutsche Gesellschaft für Soziologie</i>
DJI	German Youth Institute – <i>Deutsches Jugendinstitut</i>
<b>E</b>	
ECHP	European Community Household Panel – <i>Europäisches Haushaltspanel</i>
ELSA	English Longitudinal Study of Ageing
ESAC	European Statistical Advisory Committee
ESDS	Economic and Social Data Service
ESFRI	European Strategy Forum on Research Infrastructures
ESOMAR	European Society for Opinion and Marketing Research
ESRC	Economic and Social Research Council
ESS	European Social Survey
EU-LFS	European Labour Force Survey
EU-SILC	European Statistics on Income and Living Conditions
Eurostat	Statistical Office of the European Union

EVS	Income and Consumption Survey* – <i>Einkommens- und Verbrauchsstichprobe</i>
-----	--

## F

FiDASt	Firm-Level Data from Official Statistics* – FirmenDaten aus der Amtlichen Statistik
--------	---

## G

GEOSS	Global Earth Observation System of Systems
GESIS	Leibniz Institute for the Social Sciences – <i>Leibniz-Institut für Sozialwissenschaften</i>
GfK	Society for Consumer Research – <i>Gesellschaft für Konsumforschung</i>
GGS	Generations and Gender Survey
GIS	Geographical Information System
GLES	German Longitudinal Election Study
GLHS	German Life History Study – <i>Deutsche Lebensverlaufsstudie</i>
GNES	German National Election Study

## H

HILDA	Household, Income and Labour Dynamics in Australia
HIS	Higher Education Information System – <i>Hochschul-Informationen-System</i>
HRS	Health and Retirement Study

## I

IAB	Institute for Employment Research – <i>Institut für Arbeitsmarkt- und Berufsforschung</i>
IAB-PASS	IAB panel study “Labor Market and Social Security” – <i>Panel “Arbeitsmarkt und soziale Sicherung” des IAB</i>
IABS	IAB Employment Sample – <i>IAB-Beschäftigtenstichprobe</i>
IALS	International Adult Literary Survey
ICPSR	Inter-University Consortium for Political and Social Science Research
ICT	Information and Communication Technology

---

\* Inofficial translation.

IDF	International Data Forum
IEBS	Integrated Employment Biographies Sample – <i>Stichprobe der Integrierten Erwerbsbiografien des IAB</i>
Ifo	Ifo Institute for Economic Research – <i>ifo Institut für Wirtschaftsforschung</i>
ILO	International Labour Organization – <i>Internationale Arbeitsorganisation</i>
IQB	Institute for Educational Progress – <i>Institut zur Qualitätsentwicklung im Bildungswesen</i>
ISI	International Statistical Institute
ISR	(Michigan) Institute for Social Research
ISSP	International Social Survey Programme
IZ	Specialized Information for the Social Sciences – <i>Fachinformation für die Sozialwissenschaften</i>
IZA	Institute for the Study of Labor – <i>Forschungsinstitut zur Zukunft der Arbeit</i>

## K

KombiFiD	Combined Firm Data for Germany – <i>Kombinierte Firmendaten für Deutschland</i>
KVI	German Commission on Improving the Information Infrastructure between Science and Statistics – <i>Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik</i>

## L

LAU	Aspects of Learning Prerequisites and Learning Development* – <i>Aspekte der Lernausgangslage und der Lernentwicklung</i>
LEE data	Linked Employer-Employee data – <i>kombinierte Arbeitgeber-Arbeitnehmer Daten</i>
LIAB	Linked Employer-Employee Data from the IAB – <i>Linked-Employer-Employee Daten des IAB</i>
LIS	Luxemburg Income Study
LOGIK	German Longitudinal Study on the Genesis of Individual Competencies – <i>Longitudinalstudie zur Genese individueller Kompetenzen</i>

---

\* Inofficial translation.

LSS 2005      Cross-sectional survey “Life Situation and Social Security 2005” – *Querschnittsbefragung „Lebenssituation und Soziale Sicherung 2005“*

## **M**

MCS            Millenium Cohort Study

## **N**

NCDS           National Child Development Study  
 NEPS           National Educational Panel Study  
 NLSY           National Longitudinal Survey of Youth  
 NORC           National Opinion Research Center  
 NSF            National Science Foundation  
 NUTS           Nomenclature Of Territorial Units For Statistics / Nomenclature des unites territoriales statistiques – *Systematik der Gebietseinheiten für die Statistik*

## **O**

OECD           Organisation for Economic Co-operation and Development – *Organisation für wirtschaftliche Zusammenarbeit und Entwicklung*  
 OECD/  
 SOPEMI       Continuous Reporting System on Migration / Système d'observation permanente des migrations

## **P**

pairfam        Panel Analysis of Intimate Relationships and Family Dynamics – *Beziehungs- und Familienentwicklungspanel*  
 PIAAC          Programme for the International Assessment of Adult Competencies  
 PIRLS          Progress in International Reading Literacy Study – *Internationale Grundschul-Lese Untersuchung (IGLU)*  
 PISA           Programme for International Student Assessment – *Programm zur internationalen Schülerbewertung*  
 PSID           Panel Study of Income Dynamics  
 PUF            Public Use File

## R

RatSWD	German Data Forum (formerly named as German Council for Social and Economic Data) – <i>Rat für Sozial- und Wirtschaftsdaten</i>
RV	German Pension Insurance – <i>Deutsche Rentenversicherung</i>
RWI	Rhine-Westphalian Institute for Economic Research* – <i>Rheinisch-Westfälisches Institut für Wirtschaftsforschung</i>

## S

SHARE	Survey of Health, Ageing and Retirement in Europe
SLID	Survey of Labour and Income Dynamics
SOEP	German Socio-Economic Panel – <i>Sozio-oekonomisches Panel</i>
SUF	Scientific Use File

## T

TIMSS	Trends in International Mathematics and Science Study
-------	---

## U

UKHLS	UK Household Longitudinal Study
ULME	Study on Achievement, Motivation, and Attitudes of Students at the Beginning of Vocational Education* – <i>Untersuchungen der Leistungen, Motivation und Einstellungen zu Beginn der beruflichen Ausbildung</i>
UNECE	United Nations Economic Commission for Europe

## V

VET	Vocational Education and Training
VET-LSA	Large-Scale Assessment of Vocational Education and Training
VSKT	Insurance account sample – <i>Versichertenkontenstichprobe</i>
VVL	Completed Insured Life Courses – <i>Vollendete Versichertenleben</i>

---

\* Inofficial translation.

**W**

WGL      Gottfried Wilhelm Leibniz Scientific Community – *Leibniz  
Gemeinschaft*

WR        German Council of Science and Humanities – *Wissen-  
schaftsrat*

**Z**

ZUMA     Centre for Survey Research and Methodology – *Zentrum  
für Umfragen, Methoden und Analysen*