

Wurster, Sebastian; Feldhoff, Tobias

**Schul- und Unterrichtsqualität aus der Mehrebenenperspektive. Ist die Schule oder die Klasse die relevante pädagogische Gestaltungseinheit?**

*Zeitschrift für Pädagogik 65 (2019) 1, S. 24-39*



Quellenangabe/ Reference:

Wurster, Sebastian; Feldhoff, Tobias: Schul- und Unterrichtsqualität aus der Mehrebenenperspektive. Ist die Schule oder die Klasse die relevante pädagogische Gestaltungseinheit? - In: Zeitschrift für Pädagogik 65 (2019) 1, S. 24-39 - URN: urn:nbn:de:0111-pedocs-239252 - DOI: 10.25656/01:23925

<https://nbn-resolving.org/urn:nbn:de:0111-pedocs-239252>

<https://doi.org/10.25656/01:23925>

in Kooperation mit / in cooperation with:

**BELTZ JUVENTA**

<http://www.juventa.de>

**Nutzungsbedingungen**

Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Die Nutzung stellt keine Übertragung des Eigentumsrechts an diesem Dokument dar und gilt vorbehaltlich der folgenden Einschränkungen: Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen. Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

**Terms of use**

We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. Use of this document does not include any transfer of property rights and it is conditional to the following limitations: All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

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)

Mitglied der

  
Leibniz-Gemeinschaft

# ZEITSCHRIFT FÜR PÄDAGOGIK

Heft 1

Januar/Februar 2019

## ■ *Thementeil*

**Schul- und Unterrichtsqualität am Ende der 1970er Jahre. Re-Analysen der „Drei-Länder-Studie“ von Helmut Fend**

## ■ *Allgemeiner Teil*

Zur Geschichte des Säuglingsheims. Eine vergessene Institution des bundesdeutschen Sozialstaats

Gibt es eine neue Erziehung in der Familie?  
Konturen einer Erziehung des Beratens

Kulturen der Studienorientierung? Einzelschulische und schulstrukturelle Determinanten der Studienabsicht in der Sekundarstufe II

## Inhaltsverzeichnis

*Thementeil: Schul- und Unterrichtsqualität am Ende der 1970er Jahre.  
Re-Analysen der „Drei-Länder-Studie“ von Helmut Fend*

*Tobias Feldhoff/Eckhard Klieme/Sabine Reh*

Schul- und Unterrichtsqualität am Ende der 1970er Jahre.  
Re-Analysen der „Drei-Länder-Studie“ von Helmut Fend.

Einführung in den Thementeil ..... 1

*Monika Mattes/Sabine Reh*

Entstehung und Durchführung der Gesamtschul-Studien in den 1970er Jahren.  
Monika Mattes und Sabine Reh im Zeitzeugengespräch mit Helmut Fend .....

6

*Sebastian Wurster/Tobias Feldhoff*

Schul- und Unterrichtsqualität aus der Mehrebenenperspektive: Ist die Schule  
oder die Klasse die relevante pädagogische Gestaltungseinheit? .....

24

*Brigitte Steinert/Julia Dohrmann/Christine Schmid*

Lehrerkooperation, Unterrichtsqualität und fachliche und überfachliche  
Ergebnisse von Schüler\*innen in Englisch und Mathematik.

Eine Re-Analyse der Drei-Länder-Studie von Helmut Fend .....

40

*Julia Dohrmann/Tobias Feldhoff/Brigitte Steinert/Eckhard Klieme*

Überzeugungen von Lehrkräften, Adaptivität des Unterrichts  
und Lernergebnisse im Fach Englisch .....

56

## *Allgemeiner Teil*

*Felix Berth*

Zur Geschichte des Säuglingsheims. Eine vergessene Institution  
des bundesdeutschen Sozialstaats .....

73

<i>Jutta Ecarius/Alena Berg/Ronnie Oliveras</i>	
Gibt es eine neue Erziehung in der Familie? Konturen einer Erziehung des Beratens .....	95
 <i>Claudia Schuchart</i>	
Kulturen der Studienorientierung? Einzelschulische und schulstrukturelle Determinanten der Studienabsicht in der Sekundarstufe II .....	120
 <b>Besprechungen</b>	
 <i>Walter Herzog</i>	
Matthias Proske/Kerstin Rabenstein (Hrsg.): Kompendium Qualitative Unterrichtsforschung. Unterricht beobachten – beschreiben – rekonstruieren .....	147
 <i>Johannes Drerup</i>	
Martin Eugen Gallmann: Unschärfe in der Erziehungswissenschaft. Bilanzierung einer Wissensform. ....	149
 <b>Dokumentation</b>	
Pädagogische Neuerscheinungen .....	153
Impressum .....	U3

## Table of Contents

### *Topic: Quality of Schools and Teaching at the End of the 1970s: Re-analysis of Helmut Fend's 'Drei-Länder-Studie'*

*Tobias Feldhoff/Eckhard Klieme/Sabine Reh*

Quality of Schools and Teaching at the End of the 1970s:  
Re-analysis of Helmut Fend's 'Drei-Länder-Studie' – An introduction ..... 1

*Monika Mattes/Sabine Reh*

The Development and Execution of the 'Comprehensive School Studies'  
in the 1970s: An interview with Helmut Fend ..... 6

*Sebastian Wurster/Tobias Feldhoff*

School and Teaching Quality from a Multi-level Perspective: Is the school  
or the class the relevant pedagogical unit of action? ..... 24

*Brigitte Steinert/Julia Dohrmann/Christine Schmid*

Teacher Collaboration, Instructional Quality, and Student Outcomes:  
A re-analysis of Helmut Fend's 'Drei-Länder-Studie' ..... 40

*Julia Dohrmann/Tobias Feldhoff/Brigitte Steinert/Eckhard Klieme*

Teachers' Beliefs, Adaptivity of Teaching and Learning Outcomes in English  
as a School Subject ..... 56

### *Articles*

*Felix Berth*

Children under Three in Residential Care: Forgotten institutions  
in the early Federal Republic of Germany ..... 73

*Jutta Ecarius/Alena Berg/Ronnie Oliveras*

Is there a New Education in Families? Contours of an education  
of counseling ..... 95

*Claudia Schuchart*

The Importance of the Institutional Context: Individual school characteristics  
and upper secondary school tracks as determinants of the intention to study  
at upper secondary level ..... 120

Book Reviews .....	147
New Books .....	153
Impressum .....	U3

Sebastian Wurster/Tobias Feldhoff

# Schul- und Unterrichtsqualität aus der Mehrebenenperspektive

*Ist die Schule oder die Klasse die relevante  
pädagogische Gestaltungseinheit?*

**Zusammenfassung:** Ausgehend von Fends Konzept der Einzelschule als pädagogischer Handlungseinheit wurde untersucht, welche Bedeutung die Schule für die Erzeugung von Schul- und Unterrichtsqualität im Vergleich zu den anderen Ebenen innerhalb des Mehrebenensystems des Schulwesens hat. Analysiert wurde die Varianzverteilung zentraler Merkmale der Qualität von Schule und Unterricht sowie der Lernergebnisse zwischen den Ebenen des Schulsystems mittels hierarchischer Varianzzerlegungen. Die Befunde zeigen, dass es nicht *die* eine zentrale pädagogische Handlungseinheit (weder Klasse, noch Schule oder System) gibt. Je nach Merkmal sind unterschiedliche Ebenen bedeutsam.

**Schlagworte:** Schulqualität, Unterrichtsqualität, Lernergebnisse, Schulsystem, Mehrebenensystem

## 1. Einleitung

Als Ausgangspunkt für die Gestaltung von Schulsystemen ist die Frage nach der Bedeutung einzelner Ebenen (System, Schule, Klasse, Individuum) für die Erzeugung von Schul- und Unterrichtsqualität im Mehrebenensystem des Schulwesens zentral. Forschung im deutschsprachigen Raum zur Relevanz einzelner Ebenen geht auf Helmut Fend zurück. Die Ergebnisse der Drei-Länder-Studie von Fend (1982) zur Qualität von Gesamtschulen im Vergleich zum traditionellen Schulsystem (bestehend aus Gymnasien, Haupt- und Realschulen), hatten maßgeblichen Einfluss auf die deutsche Schulforschung (vgl. Messner, 2016). Fends Befund, „*daß die Variation der Leistungen zwischen Schulen innerhalb der einzelnen Schulformen des traditionellen Schulsystems und der Gesamtschulen sehr groß ist und in hohem Maße die Variation zwischen Schulsystemen übersteigt*“ (Fend, 1982, S. 289), war Ausgangspunkt für die Entwicklung seines Konzepts der Einzelschule als pädagogischer Handlungseinheit. Fends Befunde fielen zur damaligen Zeit in Forschung, Administration, Politik und Öffentlichkeit auf einen fruchtbaren Nährboden. Die Schulstrukturdebatten der 1960er und 1970er Jahre rund um die Einführung der Gesamtschule führten zu immer stärkeren ideologischen „Grabenkämpfen“ (vgl. Mattes, 2017). Die Hinwendung zur Einzelschule wurde von vielen als befreiender Ausweg aus der festgefahrenen Diskussion gesehen (vgl. Messner, 2016). Fend hatte maßgeblichen Anteil am Perspektivwechsel in Politik und Forschung vom Gesamtsystem hin zur Einzelschule zu Beginn der 1980er Jahre.

Die Frage, welche Bedeutung die Schule im Vergleich zu den anderen Ebenen innerhalb des Mehrebenensystems des Schulwesens für die Erzeugung von Schul- und Unterrichtsqualität hat, ist auch heute noch eine zentrale Fragestellung der Schulqualitäts- und Schuleffektivitätsforschung (vgl. Ditton, 2017). Vor dem Hintergrund der Qualitätssicherung und -entwicklung sowie der Implementation von Innovationen in Schulen ist es bedeutsam, die Relevanz der Ebenen des Schulsystems für die Schul- und Unterrichtsqualität zu kennen. Neuere Large-Scale-Studien (z. B. PISA 2009, vgl. Steinert, Hochweber & Hertel, 2013) zeigen, dass der Varianzanteil von Unterrichtsmerkmalen auf Schulebene geringer ist als auf Klassen- und Individualebene. Deshalb könnte Fend aus heutiger Perspektive die Klassenebene unter- und die Schulebene überschätzt haben. Aufgrund der Aktualität der Frage nach der Rolle der Schulebene und der Bedeutung von Fends Erkenntnissen für Forschung und Schulpraxis wird dieser Fragestellung in Form von Re-Analysen der Drei-Länder-Studie nachgegangen. Zunächst wird die Genese des Konzepts der Einzelschule als pädagogischer Handlungseinheit und daran anschließend die Bedeutung dieser im Mehrebenensystem des Schulwesens thematisiert. Als Vergleichsperspektive für die Re-Analysen wird dann der Forschungsstand zur Varianzverteilung zentraler Merkmale der Schul- und Unterrichtsqualität sowie der Lernergebnisse berichtet, bevor die Informationen und Befunde zur vorliegenden Studie beschrieben werden.

## 2. Genese des Konzepts der Einzelschule als pädagogischer Handlungseinheit im Kontext der Schulqualität

Die Entstehung des Konzepts der Einzelschule als pädagogischer Handlungseinheit und dessen Beitrag für die Schulqualitätsforschung ist nur im historischen Kontext der 1960er und 1970er Jahre zu verstehen (vgl. Messner, 2016). Diese Zeit war durch zahlreiche Bildungsreformen und einen Paradigmenwechsel in der Erziehungswissenschaft („realistische Wendung“, vgl. Roth, 1964) geprägt. Die Reformeuphorie war mit einem „Planungs-, Forschungs- und Beratungsbedarf“ (Rolff & Tillmann, 1980, S. 283) verbunden, der zahlreiche empirische Studien zur Folge hatte. Exemplarisch für diese Zeit steht die Drei-Länder-Studie (vgl. Fend, 1982).

Das damals dominante Grundverständnis war, dass Merkmale der Qualität von Schule und Unterricht primär durch die Gestaltung auf Schulsystemebene (bspw. Schulformen und deren Curricula) beeinflusst werden (vgl. Messner, 2016). Untermuert wurde diese Ansicht u. a. durch Befunde angloamerikanischer Studien, wie dem Coleman Report (vgl. Coleman et al., 1966) und dessen Fazit *that schools do not matter*. Dementsprechend zielten Reformen vor allem auf eine Veränderung schulischer Rahmenbedingungen auf der Systemebene ab, wie bspw. durch die Einführung von Gesamtschulen.

Fends zentrale Analysestrategie in der Drei-Länder-Studie zur Untersuchung der Gesamtschulen in Hessen, Nordrhein-Westfalen und Niedersachsen war deshalb ein Vergleich des traditionellen Schulsystems (Gymnasien, Realschulen und Hauptschulen) mit den neugegründeten Gesamtschulen (vgl. Fend, 1982). Er verstand die Gesamtschu-

len nicht als Schulform, zu der sie später wurde, sondern als konkurrierendes Schulsystem. Der Fokus lag somit auf einem Vergleich der Variation zwischen Schulen innerhalb der Schulformen mit der Variation zwischen den Schulsystemen (traditionelles System vs. Gesamtschulsystem). Analysiert wurden die allgemeine kognitive Grundfähigkeit (Intelligenz) und Leistungen der Schülerschaft in Jahrgangsstufe 6 und 9 in den Fächern Deutsch, Mathematik und Englisch sowie Physik (nur 9. Jahrgang). Mittels mehrfaktorieller Varianzanalysen stellte Fend fest, dass die Variation zwischen Schulen innerhalb der gleichen Schulform mit 10–18% in Englisch und 9–37% in Mathematik in der 9. Jahrgangsstufe deutlich größer ist als die zwischen traditionellem Schulsystem und Gesamtschule (2% in Englisch und 0.4% in Mathematik; Fend, 1982, S. 289).

Fend entwickelte ausgehend von diesen Befunden und seinen Arbeiten zum Schulklima (vgl. Fend, 1977) seine These der Schule als pädagogischer Handlungseinheit. Unterstützung erhielt er durch die Studie von Rutter, Maughan, Mortimore und Ouston (1979), die in Reaktion zum Coleman Report (1966) unter statistischer Kontrolle der sozialen Herkunft der Schülerschaft Unterschiede zwischen Schulen feststellte. Fend charakterisierte Schulen entsprechend als *„gemeinschaftliche Problemlösungszusammenhänge, Versuche der gemeinsamen Lösung von Aufgaben des Lehrens und der Verständigung mit der heranwachsenden Generation [...] mit den vorhandenen Ressourcen einer Schule“* (Fend, 1986, S. 275).

Für die Überprüfung der These nutzte er die damals verbreitete Methode des Extremgruppenvergleichs. Seine Fragestellung lautete: Nach welchen Merkmalen lassen sich gute und schlechte Schulen voneinander unterscheiden? Dabei identifizierte er 15 Qualitätsmerkmale (z. B. Kollegialität, Schüler-Lehrer-Beziehung; Fend, 1986, S. 275). Neben der heutigen Kritik an Extremgruppenvergleichen (vgl. Creemers, Kyriakides & Sammons, 2010) und der fehlenden Berücksichtigung der Mehrebenenstruktur, sieht Messner (2016) auch aus damaliger Perspektive Schwächen an Fends Vorgehen. Diese beziehen sich auf die Stichprobe, fehlende Informationen zu den Schulformen und teilweise unklare Kriterien für die Zuordnung zu guten oder schlechten Schulen.

Fend präsentierte 1988 eine umfassende Überarbeitung des Konzepts (vgl. Fend, 1988). Die bedeutsamste Weiterentwicklung lag in der mehrebenenanalytischen Beschreibung von Schulqualität. Als eine Art Fazit seiner Studien ordnet er den vier Ebenen Schulsystem, Schule, Klasse und Person (Lehrpersonen bzw. Schülerschaft) 35 Merkmale zu, die die Qualität auf der jeweiligen Ebene bestimmen (vgl. Abschnitt 3). Leider lässt sich nicht nachvollziehen, wie Fend diese Merkmale identifizierte. Mit der mehrebenenanalytischen Konzeption war Fend seiner Zeit jedoch voraus. Zehn Jahre später resümiert Fend (1998) nach erneuter differenzierter Analyse der Daten (vor allem in Form korrelativer und varianzanalytischer Extremgruppenvergleiche) seiner früheren Studien, dass der alleinige Fokus auf die Schulebene unangebracht ist. Der Unterricht und Kontext des Bildungssystems spielen ebenfalls eine wichtige Rolle für die Schulqualität. Fend hat damit lange vor den bedeutsamen Ergebnissen der PISA-Studien die Rolle des Unterrichts und der Lehrperson erkannt (vgl. Messner, 2016). Fend, so Messner (2016, S. 107), *„sieht also die Funktion der Schule als Handlungseinheit [...] nicht länger in ihrem statisch-qualitätsdeterminierenden Charakter, sondern in ihrer dyna-*

*mischen Bedeutung als Zentrum der schulkulturellen Praxis der Selbstreflexion, mit deren Hilfe ‚lernende Kollegien‘ die Entwicklung der Schule wesentlich vorantreiben können“.*

### **3. Bedeutung der Schule als pädagogische Handlungseinheit im Mehrebenensystem Schule**

Die Überlegungen aus Fends Konzept der Einzelschule als pädagogischer Handlungseinheit zur Erzeugung schulischer Qualität finden sich auch in der gegenwärtigen Schulqualitäts- und Schuleffektivitätsforschung wieder. So verweist bspw. Ditton (2017) auf die Bedeutung der Schulebene für die Erzeugung von Schul- und Unterrichtsqualität sowie für die Lernergebnisse. Mittlerweile hat sich die Unterscheidung von den i. d. R. vier Ebenen Individuum, Klasse, Schule und Kontext bzw. System, ähnlich wie Fend dies schon 1988 konzipierte, und deren Wechselwirkung etabliert (vgl. Creemers & Kyriakides, 2008). Auf der Individualebene der Schülerinnen und Schüler werden vor allem die Lernergebnisse in den Blick genommen. Dazu gehören kognitive Lernziele, motivational-affektive, metakognitive und verhaltensbezogene Aspekte (vgl. Creemers & Kyriakides, 2008). Auf Klassen- bzw. Unterrichtsebene haben sich als fachübergreifende Merkmale der Unterrichtsqualität die Basisdimensionen guten Unterrichts (Klassenführung, kognitive Aktivierung und ein schülerorientiertes Unterrichtsklima) etabliert (vgl. Klieme & Rakoczy, 2008). Zudem gilt die adaptive Gestaltung des Unterrichts anhand individueller Lernvoraussetzungen der Schülerschaft als wichtiges Merkmal der Unterrichtsqualität (vgl. Helmke, 2010). Im Unterschied zu den proximalen Faktoren auf der Unterrichtsebene, die einen direkten Einfluss auf die Lernergebnisse der Schülerinnen und Schüler haben, ist die verlässliche Identifizierung distaler Faktoren auf der Schulebene schwieriger. In Bezug auf zentrale Schulqualitätsmerkmale besteht deshalb weniger Einigkeit (vgl. Ditton, 2017). In der Regel werden Faktoren wie Kultur und Ethos der Schule, Führung, Kooperation sowie Personalentwicklung und Qualitätssicherung genannt. Zur System- bzw. Kontextebene werden die Aspekte sozialregionaler Kontext, schulstrukturelle Vorgaben, Ressourcen und Unterstützungsstrukturen sowie Maßnahmen der Qualitätssicherung auf Systemebene gezählt. Die jeweils höheren Ebenen stellen dabei den Handlungsrahmen für die unteren Ebenen dar.

Diese gemäß dem aktuellen Forschungsstand bedeutsamen Faktoren weisen deutliche Überschneidungen zu den in der Drei-Länder-Studie von Fend (1982) untersuchten Faktoren auf. Auf der Individualebene hat Fend schon damals nicht nur kognitive Ergebnisse in Form der Schülerleistung, sondern auch nicht-kognitive Merkmale, wie z. B. das Wohlbefinden in der Schule und im Unterricht erhoben. Auf Unterrichtsebene wurden Facetten aus zwei der drei Basisdimensionen guten Unterrichts erhoben: Merkmale zur Klassenführung (z. B. Strukturiertheit und effektive Nutzung der Lernzeit) sowie zum schülerorientierten Unterrichtsklima die Unterstützung durch die Lehrperson. Zudem wurde mit der Adaptivität des Unterrichts das Ausmaß der inneren Differenzierung erfasst. Auf der Schulebene hat Fend als Merkmale der Schulkultur und des

Ethos u. a. die Reformbereitschaft, Leistungserwartungen an die Schülerschaft, Förderorientierung, die Lehrkooperation sowie Fort- und Weiterbildung als Teil der Personalentwicklung erfasst. Lediglich das Schulmanagement und die Qualitätssicherung wurden nicht in den Blick genommen.

#### **4. Varianzverteilung zentraler Merkmale der Schul- und Unterrichtsqualität sowie der Lernergebnisse**

Die Frage, welche Bedeutung die Schule im Vergleich zu den anderen Ebenen innerhalb des Schulsystems für die Erzeugung der Schul- und Unterrichtsqualität hat, ist zentral für die Schulqualitätsforschung. Im Unterschied zu den 1970er Jahren wird dieser Frage heute mit Hilfe von Mehrebenenanalysen und der Berechnung von Intraklassenkorrelationen (ICC) nachgegangen.

Wenger, Lüdtke und Brunner (2018) analysierten auf Basis der internationalen PISA-Daten der Jahre 2000 bis 2012 aus 81 Ländern die Variabilität der Schülerurteile zu den drei Basisdimensionen guten Unterrichts. Es zeigt sich, dass sich die Schülerurteile zwischen Schulen substanziell unterscheiden, was zunächst für Fends These spricht. Aufgrund der Datenlage konnte die Klassenebene nicht separat berücksichtigt werden. Idealerweise wird zwischen der Schüler-, Klassen- und Schulebene unterschieden, da insbesondere bei Unterrichtsmerkmalen die Klasse eine zentrale Rolle spielt. Creemers und Kyriakides (2009) kommen zu dem Ergebnis, das 75 % der Varianz auf Schüler-, 15 % auf Klassen- und 10 % auf die Schulebene verteilt sind.

Für Deutschland liegen nur wenige Ergebnisse zur Varianzzerlegung von Merkmalen der Unterrichtsqualität aus repräsentativen Studien vor, die die Individual-, Klassen-, Schul- und Kontextebene unterscheiden. Entsprechende Befunde stammen aus den Studien „PISA 2009 – Nationale Ergänzungsstudie“ und „Deutsch Englisch Schülerleistungen International (DESI)“. Tabelle 1 zeigt die in Publikationen zu diesen Studien berichteten Varianzanteile auf den unterschiedlichen Ebenen für verschiedene Unterrichtsmerkmale bzw. die Schülerleistung.

Hinsichtlich der fachlichen Leistung zeigt sich ein hoher Varianzanteil auf Schulebene, der vor allem der Tatsache geschuldet ist, dass die Varianz zwischen den Schulformen inkludiert ist. Wie zu erwarten, ist der Varianzanteil auf Individualebene ebenfalls hoch. Hier fließen die unterschiedlichen Voraussetzungen und Fähigkeiten der Schülerinnen und Schüler mit ein. Bei den Unterrichtsmerkmalen zeigt sich ebenfalls der größte Teil der Varianz auf Schülerebene. In diesem Fall ist dies auf eine interindividuell unterschiedliche Wahrnehmung des Unterrichts zurückzuführen. Die Variation auf Klassenebene ist in der Regel höher (zwischen 7 und 16 %) als auf der Schulebene inklusive Schulform (zwischen 3 und 8 %).

In den bisher beschriebenen Analysen wurde die Ebene der Schulform (Kontext) nicht explizit modelliert. Vor dem Hintergrund der Genese von Fends These der Schule als pädagogischer Handlungseinheit ist es jedoch sinnvoll, möglichst alle vier Ebenen zu unterscheiden. Berücksichtigt man die Ebene der Schulform, zeigt sich hinsichtlich

	Studie	Schule/inkl. Schulform	Klasse	Schüler/innen
<i>Unterrichtsprozesse</i>				
Unterstützung	PISA 2009/Deutsch	8	16	76
	DESI/Englisch	5	10	85
	DESI/Deutsch	4	13	83
Anforderungen	DESI/Englisch	8	7	85
	DESI/Deutsch	3	12	85
Strukturiertheit	PISA 2009/Deutsch	6	10	84
	DESI/Englisch	5	9	86
	DESI/Deutsch	3	11	86
<i>Lernergebnisse</i>				
Schülerleistung (unter Kontrolle Schulform)	PISA 2009/Deutsch Lesen	15	6	79
Schülerleistung (ohne Kontrolle Schulform)		<b>Schulform</b>	<b>Schule</b>	<b>Schüler/innen</b>
	PISA 2009/Deutsch Lesen	50	3	47
	DESI/Englisch	55	9	36
	DESI/Deutsch	54	12	34
	DESI/Englisch	64	10	26
	DESI/Deutsch	52	11	37

Anmerkungen: PISA 2009: Steinert, Hochweber & Hertel (2013); DESI: Klieme, Steinert & Hochweber (2010); Steinert, Hartig & Klieme (2008).

*Tab. 1: Varianzverteilung von Unterrichtsqualitätsmerkmalen und Schülerleistungen auf die Ebenen des Schulsystems in %*

der Schülerleistung in der DESI-Studie, dass erwartungsgemäß aufgrund des Trackings durch die Schulform der größte Teil der Varianz erklärt wird, gefolgt von der Individual-ebene und der Schulebene (vgl. Tab. 1). Hinsichtlich der Varianzverteilung von Unterrichtsprozessen auf den vier Ebenen liegen Ergebnisse aus großflächigen Erhebungen der Schulinspektionen vor (vgl. Wurster & Gärtner, 2013). Bewertungen des Unterrichts durch die Schülerschaft zeigen durchschnittlich 9% Varianzanteil auf Ebene der Schulform, 6% auf Schul-, 8% auf Klassen- und 78% auf Individualebene.

## 5. Forschungsfragen

Aktuelle Large-Scale-Studien zeigen, dass der Varianzanteil von Prozessmerkmalen des Unterrichts auf Schulebene geringer ist als auf Klassen- und Individualebene. Deshalb könnte Fend in seinen damaligen Analysen aus heutiger Perspektive die Klassenebene unter- und die Schulebene überschätzt haben. Aufgrund der Bedeutung von Fends Studien für die Schulqualitätsforschung und der Aktualität der Frage zur Varianzverteilung auf die unterschiedlichen Ebenen, werden die Daten aus der Drei-Länder-Studie (vgl. Fend, 1982) re-analysiert. Es geht dabei nicht um eine Replikation von Fends Forschungsfragen und Ergebnissen mit neuen Methoden, sondern um eine Reformulierung der Forschungsfragen auf Basis unseres heutigen Wissenstandes und Erkenntnisinteresses:

- 1) Wie verteilt sich die Varianz zentraler Prozessmerkmale von Schule und Unterricht sowie von Ergebnismerkmalen der Schülerinnen und Schüler auf die einzelnen Ebenen (Individuum, Klasse, Schule, Schulform bzw. Schulsystem)? Unterscheidet sich die Verteilung der Varianz auf den Ebenen je nach Merkmal?
- 2) Lassen sich Unterschiede in der Varianzverteilung zwischen den Ebenen Ende der 1970er Jahre und heute feststellen?

## 6. Methode

### 6.1 Stichprobe

Grundlage der Analysen sind die neu aufbereiteten Daten der Drei-Länder-Studie aus den Jahren 1978 in Hessen und Nordrhein-Westfalen sowie 1979 in Niedersachsen. Die Datenbasis umfasst Befragungen von Lehrpersonen ( $N = 1091$ ) sowie die Befragung und Testung von Schülerinnen und Schülern ( $N$  vgl. Tab. 4 und 5) der 9. Jahrgangsstufe aus insgesamt 61 weiterführenden Schulen (vgl. Tab. 2). In diesen wird zum Teil leistungsdifferenziert unterrichtet, weshalb die entsprechende Analyseebene Klassen und leistungsdifferenzierte Kurse enthält.

	Hauptschulen	Realschulen	Gymnasien	Gesamtschulen	Gesamt
NRW 1978	7	4	4	9	24
Hessen 1978	3	–	–	12	15
Niedersachsen 1979	8	4	4	6	22
Gesamt	18	8	8	27	61

Tab. 2: Anzahl Schulen nach Bundesland und Schulform in der Drei-Länder-Studie von Fend (1982)

## 6.2 Instrumente

Für die Analysen wurden Skalen zu zentralen Merkmalen der Schulkultur, der Unterrichts- und Ergebnisqualität ausgewählt, die Fend in seinen Studien 1978 bzw. 1979 einsetzte und deren Relevanz sich in heutigen Studien zeigt (vgl. Kap. 3). Die Skalen wurden nach heutigen Standards (Reliabilität, Trennschärfe, Eindimensionalität, Faktorladung) überprüft und ggf. angepasst. Darüber hinaus wurden einige Skalen umbenannt, damit sie mit heutigen Konstrukten vergleichbar sind. Aus der Lehrkräftebefragung wurden Merkmale, die Fend als Schulkultur bezeichnet, ausgewählt (vgl. Tab. 3). Dazu gehören Überzeugungen von Lehrpersonen sowie Merkmale der Koordination (Kooperation im Kollegium, Chaos und Verantwortungslosigkeit). Die Skalen haben mit zwei Ausnahmen durchweg eine gute Reliabilität.

Aus der Schülerbefragung wurden zur Unterrichtsqualität vier Skalen ausgewählt, die sich an den Basisdimensionen guten Unterrichts (vgl. Klieme & Rakoczy, 2008) orientieren. Sie bilden Teilaspekte der Klassenführung und des lernunterstützenden Klimas sowie der Adaptivität des Unterrichts ab (vgl. Tab. 4). Sie sind alle reliabel. Die Skalen der Schülerbefragung wurden getrennt für den Mathematik- und Englischunterricht analysiert.

Skala	Beispielitem	N Items	M	SD	N	$\alpha$
Positive Einstellung zur Förderung von schwachen Schülerinnen und Schülern <sup>1</sup>	„Die ständige Rücksichtnahme auf die schwachen Schüler beeinträchtigt erheblich meinen Unterricht.“ (umgepolt)	6	3.60	.69	966	.69
Leistungserwartung an die Schülerinnen und Schüler <sup>2</sup>	„An dieser Schule wird viel von den Schülern verlangt.“	4	3.07	.64	982	.68
Einstellung zur Notwendigkeit von Druck beim Lernen <sup>3</sup>	„Ganz ohne Zwang wird auch der beste Lehrer nichts erreichen.“	6	3.78	.93	893	.81
Reformbereitschaft <sup>3</sup>	„Vor den Anforderungen der modernen Welt müssen die bislang gewohnten Schulformen und Unterrichtsmethoden einfach versagen.“	6	3.19	.90	885	.81
Lehrerkooperation <sup>2</sup>	„In den letzten drei Monaten habe ich mehrfach mit Kollegen Unterrichtsentwürfe und Materialien ausgetauscht.“	5	3.83	.98	947	.82
Chaos und Verantwortungslosigkeit an der Schule <sup>2</sup>	„An unserer Schule geht häufig alles drunter und drüber.“	5	2.14	.77	933	.82

Anmerkungen: Quelle: Befragung Lehrpersonen; Antwortkategorien der Studien von Fend (Krüger-Haenisch, 1978):

<sup>1</sup> (1) stimmt gar nicht (2) – (3) – (4) – (5) stimmt völlig; <sup>2</sup> (1) völlig falsch (2) eher falsch als richtig (3) unentschieden, weiß nicht (4) eher richtig als falsch (5) Stimmt genau; <sup>3</sup> (1) Ich finde das sehr falsch (2) Ich finde das falsch (3) Ich finde das eher falsch als richtig (4) Ich finde das eher richtig als falsch (5) Ich finde das richtig (6) Ich finde das sehr richtig.

Tab. 3: Prozessmerkmale auf Schulebene

Skalen	Beispielitem	N Items EN/MA	M EN/MA	SD EN/MA	N EN/MA	$\alpha$
<i>Unterrichtsprozesse</i>						
Adaptivität des Unterrichts <sup>1</sup>	„Unser Englisch-/Mathematiklehrer gibt uns meist nur solche Aufgaben, die wir auch lösen können.   ... gibt uns häufig solche Aufgaben, die wir nicht lösen können.“	9	3.54/3.59	.79/.78	1508/1655	.85
Strukturiertheit des Unterrichts <sup>1</sup>	„... fasst häufig noch einmal den Stoff zusammen, damit wir ihn uns gut merken können.   ... fasst nur sehr selten den Stoff zusammen.“	10	3.61/3.68	.87/.82	1511/1653	.87
Lernunterstützung im Unterricht <sup>2</sup>	„Wenn ich bei einer Aufgabe nicht weiter weiß, hilft mir mein Lehrer.“	7	2.77/2.80	.84/.84	1493/1508	.87
<i>Lernergebnisse</i>						
Wohlfühlen in der Schule <sup>3</sup>	„Nicht allen Schülern geht es in der Schule gleich gut. Wenn Du Deine gesamte Lage in der Schule betrachtest, wie wohl fühlst Du Dich dann?“	1	6.15/6.15	1.93/1.84	1448/1539	–
Zufriedenheit mit Lehrkräften auf Schulebene <sup>4</sup>	„Ich mag die meisten Lehrer bei uns recht gern.“	4	0.68/0.68	.34/.34	1415/1541	.72
Wohlfühlen im Unterricht <sup>1</sup>	„Ich selbst fühle mich in Englisch meistens sehr wohl.   Ich selbst fühle mich in Englisch sehr unwohl.“	6/6	3.26/3.47	1.05/1.01	1496/1561	.84
Schülerleistung EN/MA	IRT-Skalierung: Eindimensionales Rasch-Modell (WLEs)	34/24	–	–	1536/1660	.88/.75

Anmerkungen: Quelle: Befragung und Leistungstest Schülerinnen und Schüler; EN = Englisch, MA = Mathematik; Antwortkategorien der Studien von Fend (Krüger-Haenisch, 1978; Lukesch, 1979): <sup>1</sup> (1) Der untere Satz ist genau richtig (2) Der untere Satz ist eher richtig (3) Meine Meinung liegt zwischen beiden Sätzen (4) Der obere Satz ist eher richtig als der untere (5) Der obere Satz ist genau richtig; <sup>2</sup> (1) Nie (2) selten (3) manchmal (4) häufig (5) sehr häufig; <sup>3</sup> (1) Ich fühle mich in der Schule sehr unwohl – (5) Ich fühle mich in der Schule weder besonders wohl noch besonders unwohl – (9) Ich fühle mich in der Schule sehr wohl; <sup>4</sup> (1) stimmt (0) stimmt nicht.

Tab. 4: Unterrichtsprozesse auf Klassen- bzw. Kursebene und Lernergebnisse

Als Lernergebnisse wurden zum einen nicht-kognitive Outcomes wie das Wohlfühlen in Schule und Unterricht erfasst (vgl. Tab. 4). Zum anderen wurden Leistungstests in Mathematik und Englisch analysiert, die mittels eines eindimensionalen Raschmodells skaliert wurden (vgl. Rost, 2004). Im Vergleich zur ursprünglichen Auswertung wurden nur raschkonforme Items mit guten Infit- und Outfit-Parametern in die Analysen aufgenommen und Mean Weighted Likelihood Estimates (WLE) geschätzt (vgl. Warm, 1989). Der Englischtest bestand größtenteils aus C-Test-Items (Wortschatz und Grammatik) und einigen Leseverstehensaufgaben (vgl. Lukesch, 1979). Der Mathematiktest beinhaltete u. a. Aufgaben aus den Bereichen Geometrie, Arithmetik und Algebra (vgl. Lukesch, 1979). Die Reliabilität der Skalen und Tests ist zufriedenstellend.

### 6.3 Analysen

Die Varianzverteilung der genannten Konstrukte zwischen den individuellen und institutionellen Ebenen des Schulsystems wird mittels hierarchischer Varianzzerlegungen untersucht. Für die Lehrpersonenbefragung werden die drei Ebenen Schulform, Schule und Lehrperson unterschieden. Für die Schülerbefragung sind es die vier Ebenen Schulform, Schule, Unterrichtsgruppe und Schülerin bzw. Schüler. Zur besseren Vergleichbarkeit mit bisherigen Befunden werden zunächst Modelle mit der Schulform als Prädiktor anstatt als eigenständige Ebene berechnet (Modell 1). Danach wird die Schulform als Ebene mitmodelliert, da diese in der Drei-Länder-Studie eine zentrale Bedeutung hatte (Modell 2). Die Rolle der Schulebene kann so gemäß der Fragestellung isoliert analysiert werden. Zudem können die Curricula und Komposition der Schülerschaft und somit die differenziellen Lern- und Entwicklungsmilieus (vgl. Baumert, Stanat & Watermann, 2006) der verschiedenen Schulformen berücksichtigt werden.

Die Berechnung der hierarchischen Varianzzerlegungen erfolgt auf Basis von Varianzkomponenten, die in Mehrebenenanalysen ermittelt wurden. Für die Dekomposition der ebenenspezifischen Varianzanteile wurde zunächst ein Nullmodell berechnet. Die dort geschätzten Varianzen sind Grundlage der Berechnung der Varianzanteile jeder Ebene (vgl. Hox, 2010). Der Varianzanteil für Ebene 1 (Individualebene) wird aus der Varianz auf dieser Ebene dividiert durch die Summe der Varianzen aller Ebenen berechnet. Der Anteil der Varianz auf Ebene 2 (Schule bzw. Unterrichtsgruppe) wird entsprechend aus der Division der Varianz auf Ebene 2 durch die Summe der Varianzen aller Ebenen ermittelt. Für die 3- und 4-Ebenenmodelle wird dieses Vorgehen entsprechend erweitert. Die Analysen wurden mit dem Paket lme4 (vgl. Bates, Mächler, Bolker & Walker, 2015) in der Software R durchgeführt. Es wurde der REML-Schätzer verwendet, der für kleinere Stichproben geeigneter ist (vgl. Hox, 2010).

## 7. Ergebnisse

Zur Analyse der Varianzverteilung schulischer Prozessmerkmale aus der Lehrkräftebefragung wurden 3-Ebenenmodelle und korrespondierende Varianzzerlegungen berechnet (vgl. Tab. 5). Die Ergebnisse zeigen, dass die größte Variation durchweg auf Ebene der Individuen angesiedelt ist. Dies ist darauf zurückzuführen, dass die Aussagen der Lehrkräfte auf deren subjektiven Einschätzungen beruhen und die Messfehler auf Individualebene inkludiert sind. Es fällt auf, dass die Einstellung zur Förderung von schwachen Schülerinnen und Schülern im Vergleich mit Abstand den geringsten Varianzanteil auf Schul- und Schulformebene aufweist. Ermittelt man zusätzlich die Varianzkomponenten auf Schulform- und Schulebene getrennt, zeigen sich diese Unterschiede zwischen den Konstrukten erneut. Der Hauptteil der Variation liegt zwar auf Individualebene, jedoch ist mit Ausnahme des Konstrukts zur Einstellung zur Förderung schwacher Schülerinnen und Schüler ebenfalls ein substantieller Anteil an Varianz (13–25%) auf Schulebene zu verorten. Auffällig ist, dass ein beträchtlicher Teil

Skala	Modell 1		Modell 2		
	Schule inkl. Schulform	Lehrer	Schulform	Schule	Lehrer
Lehrerkooperation	24.7	75.3	20.7	13.4	65.9
Leistungserwartung an die Schülerinnen und Schüler	23.4	76.6	0.0	23.1	76.9
Positive Einstellungen zur Förderung von schwachen Schülerinnen und Schülern	7.1	92.9	1.8	6.4	91.8
Chaos und Verantwortungslosigkeit an der Schule	25.8	74.2	6.4	24.3	69.3
Reformbereitschaft	23.7	76.3	9.0	19.2	71.8
Einstellung zur Notwendigkeit von Druck beim Lernen	16.1	83.9	5.3	15.3	79.4

Anmerkungen: Modell 1: 2-Ebenen Modell mit Prädiktor Schulform auf Schulebene; Modell 2: 3-Ebenen Modell ohne Prädiktoren; Range  $N$  in Modell 1 und 2:  $N$  Schulform = 4,  $N$  Schule = 59–60,  $N$  Lehrer = 885–966.

Tab. 5: Verteilung der Varianz schulischer Prozessmerkmale auf die Ebenen des Schulsystems in %

(21%) der Varianz in der Lehrerkooperation auf Schulformebene zu finden ist, während dies bei den anderen Konstrukten nur zwischen 0% und 9% liegt. Eine Varianzanalyse zur vertiefenden Untersuchung dieser Unterschiede zeigt große Differenzen zwischen den Schulformen ( $p < .001$ ,  $\eta^2 = .22$ ). Lehrkräfte in Gesamtschulen berichten am häufigsten von Kooperation ( $M = 4.16$ ), gefolgt von Hauptschulen ( $M = 3.56$ ), Gymnasien ( $M = 3.09$ ) und Realschulen ( $M = 2.86$ ). Mit Ausnahme der Unterschiede zwischen Gymnasien und Realschulen ( $p = .65$ ,  $d = .19$ ) sind alle mittels multipler Paarvergleiche mit Bonferroni-Korrektur ermittelten Unterschiede signifikant ( $p < .001$ ,  $.39 \leq d \leq 1.10$ ). Insbesondere die Unterschiede zwischen den Gesamtschulen und allen anderen Schulformen sind groß ( $p < .001$ , Range  $d = .90$  bis  $1.10$ ). Zusammenfassend lässt sich allerdings festhalten: Schulen unterscheiden sich – mit Ausnahme der Kooperation – stärker als Schulformen, was die These der Schule als pädagogischer Handlungseinheit unterstützt.

Zur Analyse der Daten der Schülerschaft wurden die vier Ebenen Schulform, Schule, Klasse sowie Schülerin und Schüler unterschieden. Für alle Unterrichtsprozesse (vgl. Tab. 6) zeigt sich auf Ebene der Schulform kaum Varianz (0.2 bis 3%). Auf Schulebene werden zwischen 5 und 10% der Varianz verortet, während dies auf Klassenebene mit 16 bis 24% ein deutlich größerer Anteil ist. Für die untersuchten Unterrichtsprozesse liegt zwar ebenfalls der mit Abstand größte Varianzanteil auf Individualebene, jedoch zeigt sich, dass insbesondere die Klassenebene – gerade im Vergleich mit der Schulform- und Schulebene – einen substanziellen Anteil der Variation enthält. Ein ähnliches Muster findet sich für die nicht-kognitiven Lernergebnisse. Auch hier weist die Schulform nahezu keine (0 bis 1.1%), die Schulebene erwartungskonform vergleichsweise

	Fach	Modell 1			Modell 2			
		Schule inkl. Schulform	Klasse/ Kurs	Schüler/ innen	Schulform	Schule	Klasse/ Kurs	Schüler/ innen
<i>Unterrichtsprozesse</i>								
Unterstützung	EN	5.1	17.7	77.2	3.0	5.0	16.9	75.1
	MA	10.5	16.7	72.8	1.6	9.9	16.3	72.2
Strukturiertheit	EN	6.5	24.9	68.6	1.4	6.2	24.3	68.1
	MA	8.0	18.7	73.3	0.8	8.0	18.4	72.7
Adaptivität	EN	6.5	19.6	73.9	1.7	6.0	19.0	73.3
	MA	8.2	16.9	74.9	0.2	8.3	16.7	74.8
<i>Lernergebnisse</i>								
Wohlfühlen in der Schule	EN	3.0	5.2	91.8	1.1	3.0	5.4	90.5
	MA	3.3	6.3	90.5	0.0	3.3	6.5	90.2
Zufriedenheit mit Lehrkräften	EN	5.9	10.1	84.0	0.5	5.7	10.1	83.6
	MA	6.0	8.7	85.3	0.2	6.2	8.8	84.8
Wohlfühlen im Unterricht	EN	4.0	15.1	80.8	0.3	4.1	15.0	80.5
	MA	6.4	14.0	79.6	0.4	6.9	13.7	79.0
Schülerleistung	EN	24.5	34.3	41.2	45.2	0.9	22.3	31.5
	MA	22.3	22.1	55.6	30.8	5.5	17.8	45.9

Anmerkungen: Modell 1: 2-Ebenen-Modell mit Prädiktor Schulform auf Schulebene; Modell 2: 3-Ebenen-Modell ohne Prädiktoren; Range  $N$  in Modell 1 und 2 Englisch:  $N$  Schulform = 4,  $N$  Schule = 59–60,  $N$  Klasse/Kurs = 209–212,  $N$  Schüler/innen = 1281–1389; Range  $N$  in Modell 1 und 2 Mathematik:  $N$  Schulform = 4,  $N$  Schule = 60,  $N$  Klasse/Kurs = 216–220,  $N$  Schüler/innen = 1363–1504.

Tab.6: Verteilung der Varianz der Unterrichtsprozesse und Lernergebnisse auf die Ebenen des Schulsystems in %

wenig (3 bis 7%) und die Klassenebene (5 bis 15%) den relevantesten Varianzanteil auf – die Individualebene ausgenommen. Wohlfühlen in Schule und Unterricht sowie die Zufriedenheit mit Lehrkräften sind hochgradig individuelle Angelegenheiten, was sich z. T. im Vergleich zu den Unterrichtsprozessen deutlich höheren Varianzanteilen auf Individualebene zeigt. Für die Lernergebnisse zeigen sich erwartungsgemäß große Differenzen zwischen den Schulformen (31 bis 45% Varianzanteil), die auf die Selektion der Schülerschaft bei der Zuordnung zu den Schulformen und die damit einhergehende unterschiedliche Leistungsentwicklung bis zur Jahrgangsstufe 9 (vgl. Baumert, Stanat & Watermann, 2006) zurückzuführen sind (Englisch:  $p < .001$ ,  $\eta^2 = .33$ ; Mathematik:  $p < .001$ ,  $\eta^2 = .20$ ). Trotz gleicher Muster der Varianzkomponenten lassen sich Fächerunterschiede beschreiben. Während die Englischleistung im Vergleich mit der Mathematikleistung vor allem auf Schulform- und Klassenebene variiert, sind die Va-

rianzanteile in Mathematik auf Schul- und Schülerebene im Vergleich größer. Insgesamt zeigt sich hinsichtlich der Unterrichtsprozesse und Lernergebnisse ein im Vergleich zur Schulebene höherer Anteil der Varianz auf Klassenebene.

## 8. Fazit, Grenzen und Ausblick

Die Ergebnisse zur ersten Forschungsfrage zur Verteilung der Varianz auf die Ebenen des Schulsystems zeigen erwartungsgemäß, dass für die Mehrzahl der untersuchten Merkmale der Schulkultur (z. B. Kooperation im Kollegium, pädagogische Werte und Überzeugungen) nach der Individualebene der größte Varianzanteil auf der Schulebene liegt und somit die Schule für diese Merkmale die relevante pädagogische Gestaltungseinheit ist. Somit lässt sich Fends These für diese Merkmale bestätigen. Jedoch wird ebenfalls deutlich, dass für die meisten dieser Merkmale auch ein substantieller Anteil der Varianz zwischen den Schulformen zu finden ist. Dies gilt insbesondere für die Lehrerkooperation. Hier ist die Varianz auf Schulformebene größer als die auf Schulebene. Dies könnte auf die vergleichsweise stärker ausgeprägte Kooperation in Gesamtschulen zurückzuführen sein.

Hinsichtlich der Unterrichtsqualität ist – aus heutiger Sicht ebenso erwartungsgemäß – die Klassenebene die bedeutsame pädagogische Gestaltungseinheit. Allerdings ist ein nennenswerter Anteil an Varianz bei den Prozessmerkmalen des Unterrichts auch auf Schulebene zu finden, auf Ebene der Schulform dagegen nicht. Dieses Muster findet sich auch für die nicht-kognitiven Lernergebnisse (z. B. Wohlfühlen in Schule und Unterricht). Bei den Fachleistungen in Englisch und Mathematik liegt aufgrund des Trackings im deutschen Schulsystem die meiste Varianz auf Ebene der Schulform, gefolgt von der Klassenebene. Bei den Mathematikleistungen zeigt sich im Vergleich zu den Englischleistungen zudem ein nennenswerter Anteil der Varianz auf der Schulebene.

Vergleicht man diese Befunde mit Fends Ergebnissen, nach denen die Variation zwischen Schulen innerhalb der gleichen Schulform größer ist als die Variation zwischen den Schulsystemen (vgl. Kap. 2), wirkt es zunächst so, als stünden die Ergebnisse im Widerspruch zueinander. Berücksichtigt man jedoch die unterschiedlichen Herangehensweisen zur Berechnung der Varianzanteile und die dahinterstehenden Annahmen, lässt sich dieser Widerspruch auflösen. Fend verstand die Gesamtschulen als alternatives Schulsystem zum traditionellen Schulsystem (Haupt-, und Realschulen, Gymnasien). Dementsprechend war dieser Systemvergleich sein Analysefokus, während heute die Gesamtschule als eine von mehreren Schulformen gilt. Zudem hat Fend die Varianzanteile nicht auf der Klassenebene modelliert, was die größeren Varianzanteile auf Schulebene erklärt.

Zur Beantwortung der zweiten Forschungsfrage werden die re-analysierten Daten aus der Drei-Länder-Studie (vgl. Tab. 5 und 6) mit Befunden jüngerer Schulleistungstudien (PISA 2009 und DESI, vgl. Tab. 1) verglichen. Merkmale der Unterrichtsqualität und die Fachleistungen (insb. im Fach Englisch) weisen in den 1970er Jahren einen höheren Varianzanteil auf Klassenebene auf als heute. Eine mögliche Erklärung bezüg-

lich der Fachleistungen könnte in der Überrepräsentation von Gesamtschulen in der Drei-Länder-Studie liegen. Die Teilstichprobe aus Hessen besteht fast vollständig aus Gesamtschulen des sogenannten „Flächenversuchs Wetzlar“ (vgl. Mattes, 2017). Die Gesamtschulen und ein Teil der Hauptschulen arbeiteten mit einer äußeren Fachleistungsdifferenzierung, bei der in einer Jahrgangsstufe bis zu drei Kurse pro Fach auf unterschiedlichen Leistungsniveaus existierten. Möglicherweise ist die größere Varianz auf den hohen Anteil an Gesamtschulen in der Drei-Länder-Studie zurückzuführen.

Zusammenfassend wird deutlich, dass es nicht *die eine* zentrale pädagogische Handlungseinheit (weder Klasse, noch Schule oder System) gibt, sondern, dass je nach Merkmal unterschiedliche Ebenen bedeutsam sind bzw. Schul- und Unterrichtsqualität durch die Interaktion der verschiedenen Merkmale und über die verschiedenen Ebenen hinweg entsteht (vgl. bspw. Creemers & Kyriakides, 2008). Somit war Fends Weiterentwicklung seines Konzepts der Schule als pädagogischer Handlungseinheit hin zu einem mehrebenenanalytischen Verständnis von Schulqualität folgerichtig und wegweisend. Zudem zeigen die Befunde, dass sich die Varianzanteile je nach Merkmal und teilweise zwischen den Fächern sehr unterschiedlich auf die drei bzw. vier Ebenen verteilen. Somit lohnen sich fachspezifisch differenzierte Mehrebenen-Varianzanalysen, die als Ausgangspunkt für vertiefende Mehrebenenregressionsanalysen dienen können (siehe folgende Beiträge in diesem Heft).

Die Drei-Länder-Studie von Fend (1982) weist ähnliche Grenzen wie gegenwärtige Forschung auf. Die Daten stellen keine Zufallsstichprobe dar. Zur Analyse möglicher Veränderungen von Schulqualitätsmerkmalen im Zeitvergleich wird ein Projekt zum Vergleich der Befunde der Drei-Länder-Studie mit aktuellen Daten aus einer Follow-Up-Studie an den gleichen Schulen durchgeführt. Ergänzend können Sekundäranalysen von Large-Scale-Assessments Auskunft über die aktuelle Varianzverteilung der Fachleistungen und Prozesse auf Schul- und Unterrichtsebene geben. Zudem ist die Frage nach Faktoren, die die Varianz auf den einzelnen Ebenen erklären, zu analysieren. Nichtsdestotrotz haben die Re-Analysen der Drei-Länder-Studie gezeigt, dass je nach Schulqualitätsmerkmal sowohl die Schul- wie auch Unterrichtsebene relevante Gestaltungseinheiten sind.

## Literatur

- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1–48.
- Baumert, J., Stanat, P., & Watermann, R. (Hrsg.) (2006). *Herkunftsbedingte Disparitäten im Bildungswesen: Differenzielle Bildungsprozesse und Probleme der Verteilungsgerechtigkeit*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., & Weinfeld, F. D. (1966). *Equality of Educational Opportunity*. Washington/DC: Government Printing Office.
- Creemers, B. P. M., & Kyriakides, L. (2008). *The Dynamics of Educational Effectiveness*. London/New York: Routledge.

- Creemers, B. P. M., & Kyriakides, L. (2009). Situational Effects of the School Factors Included in the Dynamic Model of Educational Effectiveness. *South African Journal of Education*, 29, 294–315.
- Creemers, B. P. M., Kyriakides, L., & Sammons, P. (Hrsg.) (2010). *Methodological Advances in Educational Effectiveness Research*. London/New York: Routledge.
- Ditton, H. (2017). Zum Wirkungszusammenhang der schulischen Handlungsebenen. In U. Steffens, K. Maag Merki & H. Fend (Hrsg.), *Schulgestaltung* (S. 59–84). Münster: Waxmann.
- Fend, H. (1977). *Schulklima: Soziale Einflussprozesse in der Schule*. Weinheim: Beltz.
- Fend, H. (1982). *Gesamtschule im Vergleich: Bilanz der Ergebnisse des Gesamtschulversuchs*. Weinheim: Beltz.
- Fend, H. (1986). „Gute Schulen – schlechte Schulen“. Die einzelne Schule als pädagogische Handlungseinheit. *Die Deutsche Schule*, 78(3), 275–293.
- Fend, H. (1988). Schulqualität. Die Wiederentdeckung der Schule als pädagogische Gestaltungsebene. *Neue Sammlung*, 28(4), 537–547.
- Fend, H. (1998). *Qualität im Bildungswesen*. Weinheim: Juventa.
- Helmke, A. (2010). *Unterrichtsqualität und Lehrerprofessionalität*. Seelze: Klett-Kallmeyer.
- Hox, J. J. (2010). *Multilevel Analysis: Techniques and applications* (2. Aufl.). New York: Routledge.
- Klieme, E., & Rakoczy, K. (2008). Empirische Unterrichtsforschung und Fachdidaktik: Outcome-orientierte Messung und Prozessqualität des Unterrichts. *Zeitschrift für Pädagogik*, 54(2), 222–237.
- Klieme, E., Steinert, B., & Hochweber, J. (2010). Zur Bedeutung der Schulqualität für Unterricht und Lernergebnisse. In W. Bos, E. Klieme & O. Köller (Hrsg.), *Schulische Lerngelegenheiten und Kompetenzentwicklung* (S. 231–255). Münster: Waxmann.
- Krüger-Haenisch, E.-M. (1978). *Erhebungsinstrumente der Leistungsstudie 1978*. Universität Konstanz.
- Lukesch, H. (1979). *Schulleistungsvergleiche zwischen Gesamtschulen in Hessen und Schulen des gegliederten Schulsystems am Ende der Pflichtschulzeit* (Arbeitsbericht Nr. 13). Universität Konstanz.
- Mattes, M. (2017). Von der Leistungs- zur Wohlfühlschule? Gesamtschule als Ort gesellschaftlicher Debatten und pädagogischer Wissensproduktion in der Bundesrepublik Deutschland der 1970/80er Jahre. In S. Reh, B. Behm, T. Drope & E. Glaser (Hrsg.), *Wissen machen. Beiträge zu einer Geschichte erziehungswissenschaftlichen Wissens in Deutschland zwischen 1945 und 1990* (63. Beiheft, Zeitschrift für Pädagogik, S. 187–206). Weinheim/Basel: Beltz Juventa.
- Messner, R. (2016). Die Einzelschule als pädagogische Handlungseinheit und das Zusammenspiel der Handlungsebenen und institutionellen Akteure. In U. Steffens & T. Bargel (Hrsg.), *Schulqualität – Bilanz und Perspektiven* (S. 95–114). Münster: Waxmann.
- Rolff, H.-G., & Tillmann, K.-J. (1980). Schulentwicklungsforschung: Theoretischer Rahmen und Forschungsperspektive. In H.-G. Rolff, K. Klemm, K.-J. Tillmann & K. H. Hansen (Hrsg.), *Jahrbuch der Schulentwicklung* (S. 237–265). Weinheim: Beltz.
- Rost, J. (2004). *Lehrbuch der Testtheorie – Testkonstruktion* (2. Aufl.). Bern: Huber.
- Roth, H. (1964). Die realistische Wendung in der pädagogischen Forschung. In H. Röhrs (Hrsg.), *Erziehungswissenschaft und Erziehungswirklichkeit* (S. 179–191). Frankfurt: Akademische Verlagsanstalt.
- Rutter, M., Maughan, B., Mortimore, P., & Ouston, J. (1979). *Fifteen Thousand Hours: Secondary schools and their effects on children*. Cambridge: Harvard University Press.
- Steinert, B., Hartig, J., & Klieme, E. (2008). Institutionelle Bedingungen der Sprachkompetenz: Sprachliche Kompetenzen im Deutschen und Englischen. Schularvergleich. In DESI-Konsortium (Hrsg.), *Unterricht und Kompetenzerwerb in Deutsch und Englisch* (S. 411–451). Weinheim/Basel: Beltz.

- Steinert, B., Hochweber, J., & Hertel, S. (2013). Lesekompetenz und Lesefreude von Schülerinnen und Schülern und bildungsstandardbezogene Kompetenzüberzeugungen und Lerngelegenheiten in Schule und Unterricht. In N. Jude & E. Klieme (Hrsg.), *PISA 2009 – Impulse für die Schul- und Unterrichtsforschung* (59. Beiheft, Zeitschrift für Pädagogik, S. 12–39). Weinheim/Basel: Beltz Juventa.
- Warm, T.A. (1989). Weighted Likelihood Estimation of Ability in Item Response Theory. *Psychometrika*, 54(3), 427–450.
- Wenger, M., Lüdtke, O., & Brunner, M. (2018). Übereinstimmung, Variabilität und Reliabilität von Schülerurteilen zur Unterrichtsqualität auf Schulebene. *Zeitschrift für Erziehungswissenschaft*, <https://link.springer.com/article/10.1007/s11618-018-0813-3> [13.09.2018].
- Wurster, S., & Gärtner, H. (2013). Erfassung von Bildungsprozessen im Rahmen von Schulinspektion und deren potenzieller Nutzen für die empirische Bildungsforschung. *Unterrichtswissenschaft*, 41(3), 216–234.

**Abstract:** Based on Fend's concept of the individual school as a pedagogical unit of action, the importance of school for the creation of school and teaching quality was examined in comparison to the other levels within the multi-level school system. The variance distribution of the two main factors of school quality and teaching quality, as well as the learning outcomes between the levels of the school system, were analysed by hierarchical variance decomposition. The findings show that there is not one single pedagogical unit of action (neither class nor school or system). Depending on the factor, different levels are important for the quality of school and teaching.

**Keywords:** School Quality, Teaching Quality, Learning Outcomes, School System, Multi-level System

### **Anschrift der Autor\_innen**

Dr. Sebastian Wurster, Johannes-Gutenberg-Universität Mainz,  
 Institut für Erziehungswissenschaft,  
 Jakob-Welder-Weg 12, 55128 Mainz, Deutschland  
 E-Mail: wurster@uni-mainz.de

Prof. Dr. Tobias Feldhoff, Johannes-Gutenberg-Universität Mainz,  
 Institut für Erziehungswissenschaft,  
 Jakob-Welder-Weg 12, 55128 Mainz, Deutschland  
 E-Mail: feldhoff@uni-mainz.de