

Bos, Wilfried

## Empirisch-vergleichende Forschung. Konzeptionelle Probleme und politischer Nutzen

*Tertium comparationis* 5 (1999) 2, S. 118-133



Quellenangabe/ Reference:

Bos, Wilfried: Empirisch-vergleichende Forschung. Konzeptionelle Probleme und politischer Nutzen - In: *Tertium comparationis* 5 (1999) 2, S. 118-133 - URN: urn:nbn:de:0111-opus-29423 - DOI: 10.25656/01:2942

<https://nbn-resolving.org/urn:nbn:de:0111-opus-29423>

<https://doi.org/10.25656/01:2942>

in Kooperation mit / in cooperation with:



**WAXMANN**  
[www.waxmann.com](http://www.waxmann.com)

<http://www.waxmann.com>

### 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



## Empirisch-vergleichende Forschung – Konzeptionelle Probleme und politischer Nutzen

*Wilfried Bos*

*Universität Dortmund*

### *Abstract*

In this article the relation of practice, theory and metatheory is discussed for comparative education. Subsequently examples of quantitatively orientated empirical research methods within the scope of comparative education are given. Thereby the value of empirical research methods is revealed for national studies as well as for the transcultural testing of theories. However, the author constantly points out that empirical educational research in the scope of comparative education does not replace historical hermeneutic research but to complement it.

### 1 Vorbemerkung

Dies ist die schriftliche Fassung des Vortrages gleichen Titels gehalten auf der gemeinsamen Jahrestagung der Kommissionen „Vergleichende Erziehungswissenschaft“; „Internationale kooperative Bildungsforschung“ und der AG „Interkulturelle Bildung“ in der DGfE am 12.2.1999 in Münster. Um den Vortragscharakter nicht zu verfälschen wurden weitestgehend rhetorische Figuren, Polarisierungen, Provokationen und Kontrastierungen, wie sie in mündlichen Vorträgen, erst recht gehalten zu Ehren eines Emeritus, üblich sind, beibehalten. Als ich diesen Vortrag hielt, stand ich noch unter dem Eindruck der erst einen Tag zurückliegenden „Professionspolitischen Konferenz“ der DGfE zum Thema „Erziehungswissenschaft in Studium und Beruf“ 10.2.99-11.2.99 in Dortmund, auf dem intensiv versucht wurde, über das Verhältnis Praxis, Theorie und Methoden zu diskutieren. Assoziativ verband ich diesen Versuch – vielleicht durch die Nähe zu Borussia beeinflusst – mit dem Versuch über Fußball zu diskutieren. Die Praxis ist in diesem Bild die Performance auf dem Platz, die Theorie sind die Trainingskonzepte, die Metatheorie die Reflexion und eventuelle Bewertung verschiedener Konzepte. In Dortmund viel es schwer über Praxis, über Realität, zu reden. Trefflich wurde aber über Theorie und Metatheorie gestritten: Dürfen Spielerfrauen ins Trainingslager? Welches Menschenbild steckt dahinter wenn ja bzw. nein. Welche kulturellen Traditionen werden hier berührt, welche historischen Wurzeln und Diskussionlinien müssen thematisiert werden? Welche Auswirkungen hätte dies auf die Globalisierung

und die „Dritte Welt“? Über all diese wichtigen Diskussionen und Reflexionen ließen sich Regalmeter in Bibliotheken füllen, neue Arbeitsgemeinschaften in der DGfE, evtl. neue Widmungen von Professuren, mit etwas Glück sogar einige Forderungen zur „Trainerbildung“ ableiten. Über die Reflexion der Reflexion ließen sich auf Kongressen große Debatten, unter Einbeziehung wichtiger ausländischer Gastredner, vielleicht sogar per „liveschaltung“ im Rahmen einer Videokonferenz, führen – bis man nahezu die Bodenhaftung verliert und sich wie Teile der Allgemeinen Pädagogik in die Gefahr begibt, esoterisch und überflüssig zu werden. In diesem Diskussionsumfeld befindet sich in Deutschland in der Erziehungswissenschaft der Empiriker, erst recht der Schulleistungsforscher, der mit vorsichtiger Stimme anfragt, ob man denn nicht auch mal die Tore zählen sollte.

## 2 Einleitung

Damit bin ich beim Thema. Der Titel dieses Beitrages beinhaltet eine zweifache Provokation. Einmal „Empirie“ und zum anderen „politischer Nutzen“, wobei letzterer noch nach „fachpolitischem“ und „gesellschaftlichem“ zu unterscheiden wäre. Empirie ist eine schwierige und gefährliche Angelegenheit. Schwierig, da „Handwerkzeug“ gekonnt eingesetzt werden muss. Dies setzt jahrelange Ausbildung und intensives Lernen voraus. Normative Stellungnahmen dagegen eher eine „richtige Grundhaltung“. Gefährlich ist „Empirie“, da bei solchen Untersuchungen ja etwas herauskommen kann, dass man so gar nicht haben bzw. wissen will. Darüber hinaus könnten durch Ergebnisse empirischer Forschung liebgewonnene Vorurteile angezweifelt, wenn nicht gar widerlegt werden. Von fachpolitischem Nutzen bzw. von Nutzen für das Fach ist empirische Forschung wohl unbestritten, bilden sie doch den Prüfstein für Theorie, generieren und belegen sie doch „Erkenntnis“ auf deren Basis ja erst ernsthaft gestritten bzw. reflektiert werden kann. Dabei muss man anscheinend in Kauf nehmen, dass empirische Arbeiten oft mehr Fragen aufwerfen als sie beantworten. Gesellschaftspolitisch könnte, die Betonung liegt auf „könnte“, empirisch pädagogische Forschung von Nutzen sein. „Erkenntnis“ trägt generell zur Mündigkeit zur Emanzipation aller Beteiligten bei, wenn diese „Erkenntnis“, die Fakten – hier politisches Steuerungswissen auf Grundlage empirisch pädagogischer Forschung – denn allen Beteiligten zur Verfügung steht und gleichberechtigt in einem diskursiven Prozess zur Veränderung gesellschaftlicher Wirklichkeit genutzt werden kann. Im Rahmen „Vergleichender Erziehungswissenschaft“ spielt die empirische pädagogische Forschung in Deutschland – auch auf Grund der geisteswissenschaftlichen Tradition der Pädagogik in unserem Land – eher eine Nebenrolle. Einige der Möglichkeiten vergleichender empirischer pädagogischer Forschung will ich im folgenden, an Detlef Glowkas Wort denkend, dass „abstrakte Methodendiskussion nichts bringt“, an einigen Beispielen, an denen dann punktuell konzeptionelle Probleme und evtl. politischer Nutzen aufgezeigt werden, verdeutlichen.

3 Empirie, Theorie und Praxis in vergleichender pädagogischer Forschung

Vergleichende Erziehungswissenschaft kann sich, wie Abbildung eins zeigt, in ganz unterschiedlichen Bereichen pädagogischer Forschung verorten. International vergleichend könnten z.B. „Praxisthemen“ wie Schulorganisation, Schulmanagement, Lehrerfortbildung etc. bearbeitet werden. Ob „Praxis“ in einem Land allein, dargestellt als „Länderstudie“, den Kriterien Vergleichender Pädagogischer Forschung genügt, ist strittig, findet sich aber regelmäßig in Form von Examensarbeiten oder als Einzelbeitrag in einschlägigen Zeitschriften oder Sammelbänden.

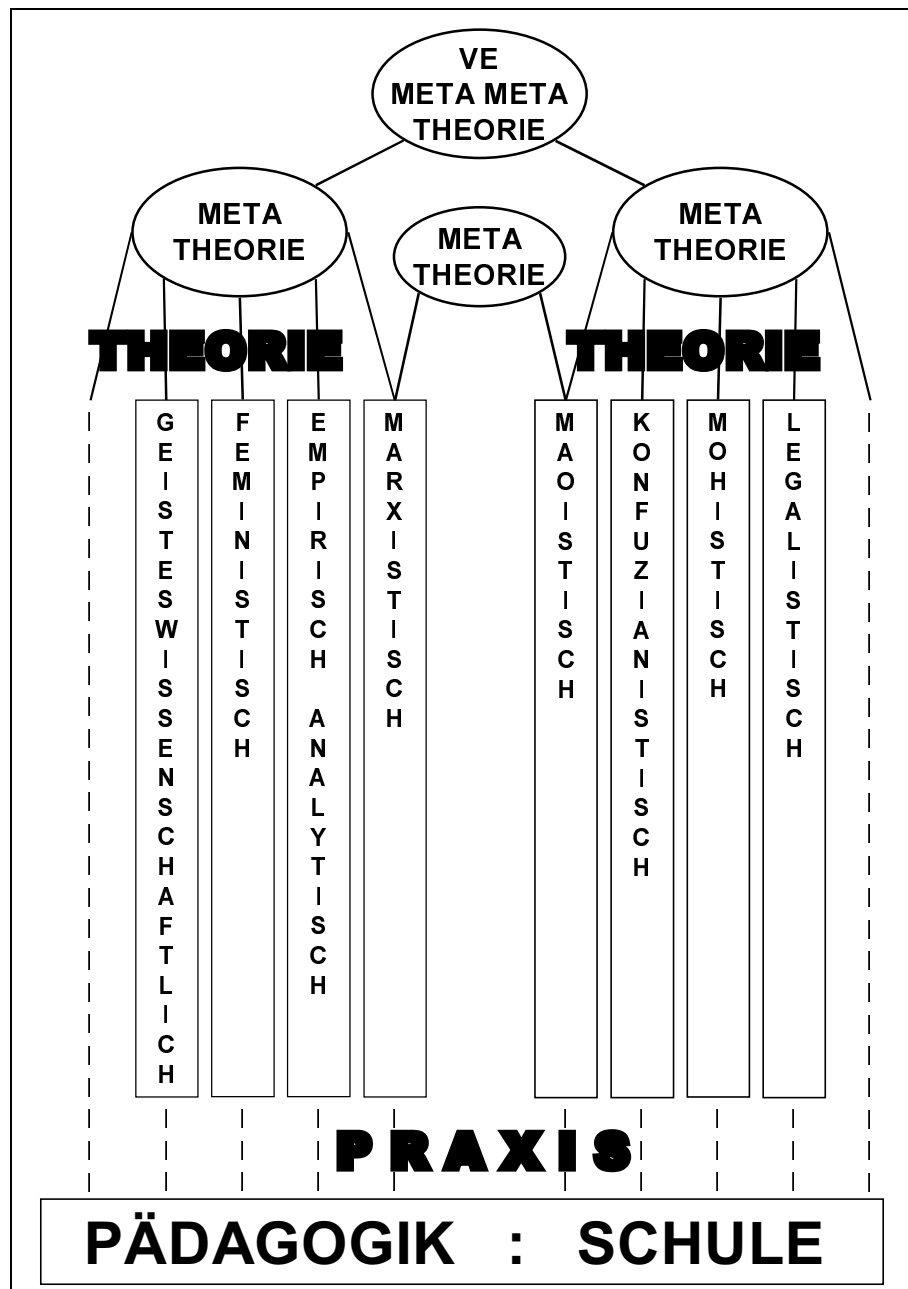


Abbildung 1: Bezugsschema zum Verhältnis von Praxis, Theorie und Empirie in der Vergleichenden Erziehungswissenschaft.

Unstrittig ist, dass die Bearbeitung einschlägiger Theoriestränge eines Landes im Rahmen einer „Metatheoriediskussion“ zu den Themen vergleichender pädagogischer Forschung gehört. In dem hier vorgelegten Bezugsschema wäre dies vielleicht der jeweilige Einfluss des Militärtheoretikers Han Fei Zi (Legalismus), des Bildungstheoretikers Meng Zi (Konfuzianismus), des Sozialutopisten Mo Zi (Mohismus) oder des modernen Agrarkommunisten Mao Zi Dong auf „Bildungsziele chinesischer Schulpädagogik von den Anfängen bis zur Gegenwart“. Die von mir gewählten „Theoriestränge“ sind dabei in keiner Weise „systematisch“, sie sollen nur die Vergleichsmöglichkeiten illustrieren. Eine andere Form metatheoretischer vergleichender pädagogischer Forschung wäre z.B. der Vergleich von „Kontroversen Bildungszielen europäischer und chinesischer Marxisten nach dem zweiten Weltkrieg“. Eine Besonderheit metatheoretischer Forschung wäre im Rahmen der vergleichenden pädagogischen Forschung der Vergleich von Metatheorien – quasi eine Meta-Metatheoriediskussion. Es können aber auch Resultate pädagogischer Praxis wie Schülerleistungen in ausgewählten Fächern international vergleichend beforscht werden. Diese internationale Schulleistungsforschung ist aber nur ein Aspekt vergleichender pädagogischer Forschung, auf den später noch etwas ausführlicher eingegangen wird.

Im folgenden möchte ich auf verschiedenen Ebenen des Vergleichs Beispiele empirischer pädagogischer Forschung vorstellen.

#### 4 Empirisch vergleichende pädagogische Forschung in Länderstudien

Im ersten Beispiel empirisch vergleichender pädagogischer Forschung wird ein inhaltsanalytisches Vorgehen im Rahmen einer „Länderstudie“ vorgestellt, eine Schulbuchanalyse von Sprachlehrbüchern der ersten Jahrgangsstufe aus der Volksrepublik China. Es wurden insgesamt 596 Schulbuchseiten aus den Jahren 1963 – vor der Kulturrevolution, 1978 – nach der Kulturrevolution und 1982 – nach dem der Reformler Deng Xiao Ping seine Position gesichert hatte, analysiert. Dabei wurde theoriegeleitet ein Kategoriensystem, wie in Abbildung zwei dargestellt entwickelt.

Der grundlegende funktionalistische Theorieansatz Parsons wurde von Fend für erziehungswissenschaftliche Fragestellungen dahingehend konkretisiert, indem er dem Erziehungswesen die gesellschaftlichen Funktionen der Qualifikation, der Integration und der Selektion zuwies. Kollege Straka und ich modifizierten diesen Ansatz für China, indem wir die Selektionsfunktion durch eine Identifikationsfunktion ersetzten, denn einerseits war unklar, inwieweit in einem kommunistischen Land gesellschaftlicher Aufstieg an schulische Leistungen gekoppelt ist und andererseits in einem Vielvölkerstaat wie China Schule und Lehrmaterialien zu einer Identifikation mit dem „neuen Staat“ beizutragen hat. Mit diesem Kategoriensystem, das durch einen iterativen Prozess zwischen Induktion und Deduktion verfeinert wurde, konnten schließlich die Buchseiten als Analyseinheit mit zufriedenstellender Interkoderreliabilität verkodet werden.

<b>Qualifikation „was“ Inhalt</b>	<b>Integration „wodurch“ Struktur</b>	<b>Identifikation „wohin“ Zielvorstellung</b>
<b>Fachliche Qualifikation (K1)</b> (Fertigkeiten, Kenntnisse um Natur, Industrie usw.)	<b>Familie (K4)</b> (Kernfamilie)	<b>China als Einheitsstaat (K7)</b> (Abgrenzung gegenüber dem Auslande, dem Nichtkommunismus)
<b>Moralische Qualifikation (K2)</b> (Tugenden, Haltungen, Sparsamkeit, Sorgfalt usw.)	<b>Schule (K5)</b> (Klassenverband, Schulkameraden)	<b>Errungenschaften der neuen Gesellschaft (K8)</b> (Modernisierung, Technik, Gesundheitswesen, Landwirtschaft usw.)
<b>Kulturelle Qualifikation (K3)</b> (Wissen um Kunst, Literatur, Geschichte usw.)	<b>Verbände und Organisationen (K6)</b> (rote Garden, junge Pioniere, KPC, Militär usw.)	<b>Sozialismus als Ideal (K9)</b> (sich für die Gemeinschaft einsetzen usw.)

Abbildung 2: Theorieorientiertes Kategoriensystem zur Inhaltsanalyse chinesischer Schulbücher.

Eine erste Häufigkeitsanalyse ergab, wie aus Tabelle eins zu ersehen ist, dass sich einige Veränderungen in der Anzahl des Vorkommens einzelner Kategorien in den verschiedenen Jahrgängen abzeichneten. Insbesondere die Kategorie neun „Sozialismus als Ideal“ verliert an relativer Bedeutung, sowohl in den Lesetexten als auch besonders in den Übungen, mit denen ja gerade auf besonders wichtige – zu lernende – Inhalte hingewiesen wird. Auf der anderen Seite wird im Laufe der Jahre die Sozialisationsinstanz „Verbände und Organisationen“ stärker in den Vordergrund gerückt. Relativ stabil über die Zeit bleibt dagegen die Position der Sozialisationsinstanz „Familie“ im unteren Drittel der Nennungshäufigkeiten.

Neben einfacher Häufigkeitsanalysen lassen sich korrelative Zusammenhänge – hier mittels des Tanimotokoeffizienten – von Texten bzw. Übungen darstellen. Damit wird dargestellt, ob in Texten und Übungen systematisch bestimmte Kategorien gemeinsam angesprochen werden, ob z.B. in Texten oder Übungen regelmäßig die Qualifikationskategorie „Fachliche Qualifikation“ und die Sozialisationsinstanz „Schule“ kombiniert werden. Die Tanimotokoeffizienten liegen hier zwischen null und hundert. Je größer der Wert, um so enger ist der Zusammenhang, der Verknüpfungsgrad. Wie in Tabelle zwei ersichtlich bleibt über alle Messzeitpunkte die „Fachliche Qualifikation“ an die

### Lehrmaterialien 1963

Lesetexte			Übungen		
Kategorie	Rang	in %	Kategorie	Rang	in %
K2	1	29	K1	1	28
K1	2	27	K2	2	26
K5	3	11	K5	3	12
K9	4	9	K9	4	9
K6	5	9	K6	5	8
K7	6	5	K4	6	6
K4	7	4	K8	7	6
K8	8	4	K7	8	4
K3	9	1	K3	9	2
Basis = 212		99	Basis = 141		101

### Lehrmaterialien 1978

Lesetexte			Übungen		
Kategorie	Rang	in %	Kategorie	Rang	in %
K1	1	27	K1	1	23
K2	2	25	K2	2	22
K9	3	14	K5	3	14
K6	4	9	K8	4	14
K5	5	8	K9	5	10
K8	6	8	K6	6	10
K4	7	4	K4	7	3
K7	8	4	K7	8	3
K3	9	3	K3	9	2
Basis = 244		102	Basis = 185		101

### Lehrmaterialien 1990

Lesetexte			Übungen		
Kategorie	Rang	in %	Kategorie	Rang	in %
K1	1	43	K1	1	30
K2	2	22	K2	2	27
K6	3	8	K5	3	14
K5	4	7	K6	4	10
K9	5	5	K3	5	5
K3	6	5	K4	6	4
K7	7	5	K8	7	4
K4	8	4	K9	8	4
K8	9	4	K7	9	3
Basis = 136		100	Basis = 92		100

Tabelle 1: Kategorienhäufigkeiten nach Jahrgang in Prozent.

Lektionstexte 1963									Übungsbereiche 1963									
	K1	K2	K3	K4	K5	K6	K7	K8		K1	K2	K3	K4	K5	K6	K7	K8	
K2	24								K2	27								
K3									K3									
K4									K4	15								
K5	23		10							K5	26							
K6	10								K6									
K7									K7	22								
K8	14							K8	17									
K9	11							K9	16		15		27		17			

Lektionstexte 1978									Übungsbereiche 1978								
	K1	K2	K3	K4	K5	K6	K7	K8		K1	K2	K3	K4	K5	K6	K7	K8
K2	13								K2	19							
K3									K3								
K4	10								K4								
K5	11							K5	13	25							
K6	12							K6	21		14		16				
K7			15		11		11		K7	13							
K8	12							K8	20	18		21		26			
K9	19							K9	13		15		14				

Lektionstexte 1982									Übungsbereiche 1982									
	K1	K2	K3	K4	K5	K6	K7	K8		K1	K2	K3	K4	K5	K6	K7	K8	
K2	16								K2	24								
K3	13								K3	16								
K4									K4									
K5	18		10							K5	15	37						
K6	15								K6	19								
K7									K7									
K8									K8	15								
K9									K9	13								

Tabelle 2: Korrelative Beziehungen zwischen den Analysekatogorien (gerundete Tanimotokoeffizienten > 10).

„Moralische Qualifikation“ gebunden. Wird vor und unmittelbar nach der Kulturrevolution die „Fachliche Qualifikation“ noch regelmäßig mit der Kategorie acht, den „Errungenschaften der neuen Gesellschaft“ in Verbindung gebracht, verliert diese Kombination in den Lehrbüchern von 1982 an Bedeutung, die wichtigere Rolle der „Schule“ als wesentliche Sozialisationsinstanz sowohl „Fachlicher“, „Moralischer“ als auch „Kultureller Qualifikation“ zeichnet sich dagegen stärker ab.



Texte und Übungen von Schulbüchern sind komplex, regelmäßig werden je Analyseeinheit mehr als zwei Kategorien miteinander kombiniert. Durch multivariate Analyseverfahren – hier mittels der Clusteranalyse – lassen sich solche systematischen Mehrfachverknüpfungen, hier als Baumdiagramm bzw. Dendrogramm dargestellt, abbilden.

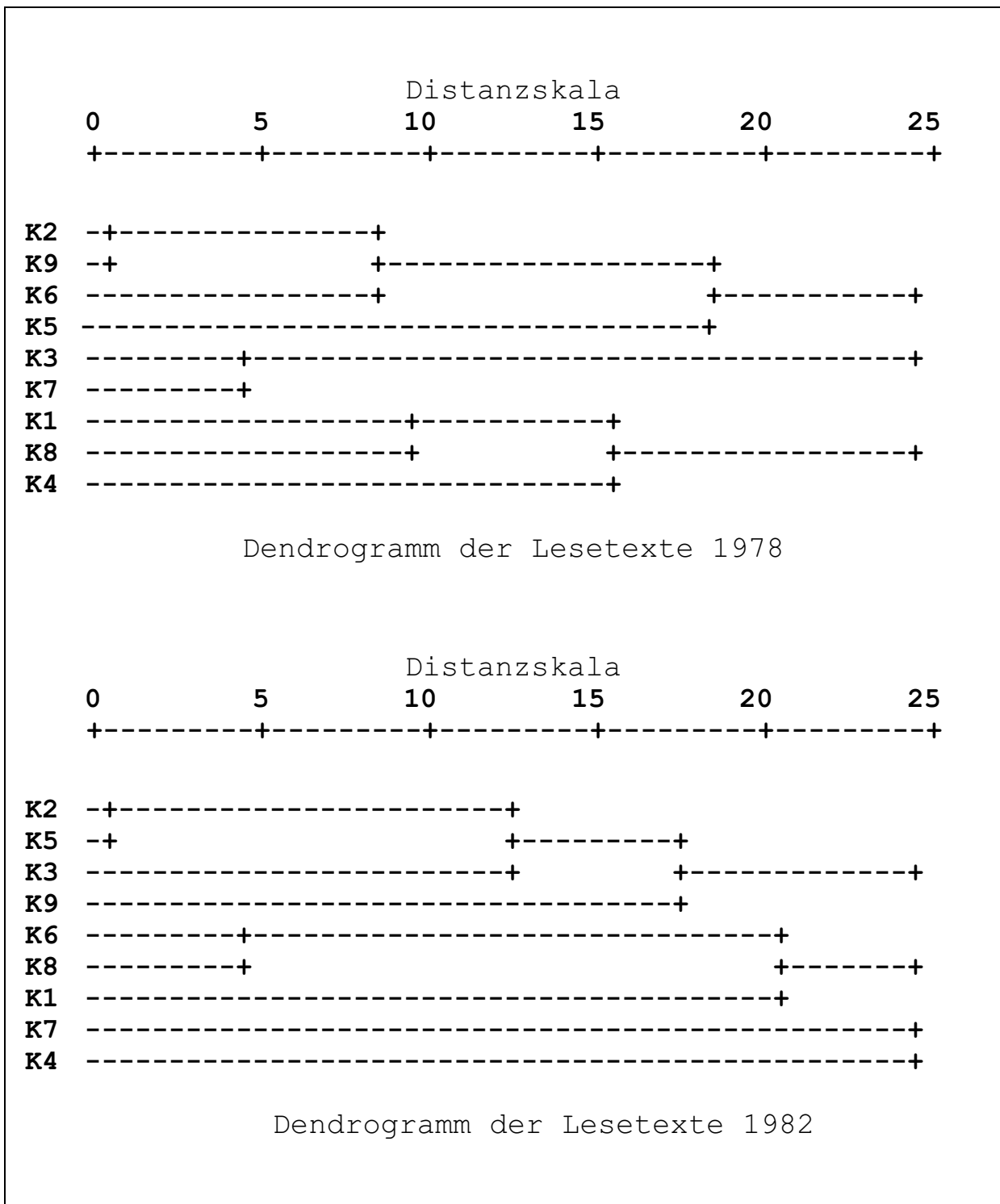


Abbildung 3: Mehrfachverknüpfungen von Analysekatogorien (Dendrogramme auf Basis von Clusteranalysen).

Die durch Kreuze miteinander verbundenen Linien der Kategorien veranschaulichen die Binnengruppierung des Kategoriensystems. Je näher die Verbindung der Kategorien auf der Distanzskala bei null liegt, um so enger ist der Zusammenhang. Ist die Sozialisationsinstanz „Familie“ 1978 noch mit der „Fachlichen Qualifikation“ und der Zielvorstellung „Errungenschaften der neuen Gesellschaft“ verbunden, steht diese Kategorie 1982 isoliert ohne Verbindung zu anderen Kategorien. Dies gilt auch für die Zielvorstellung „China als Einheitsstaat“, die in den Lesetexten von 1978 noch mit der „Kulturellen Qualifikation“ verbunden war.

Deutlich wird, hoffentlich, das auch bei „Länderstudien“ innerhalb der vergleichenden pädagogischen Forschung empirische Erhebungs- und Analyseverfahren ihren Wert haben und tiefgehende Einblicke erlauben. Allerdings kann und darf bei vergleichender empirischer Forschung die „historische und landeskundliche Dimension“ nicht vernachlässigt werden, denn sonst können eklatante Fehler auftreten. Nehmen wir hier nur als Beispiel einen Text aus einem der Schulbücher in dem berichtet wird, dass die Regentropfen zur Erde fallen, sich zu Bächlein verbinden, diese zu Flüssen und Strömen werden und schließlich im Meer zusammenkommen. Elementares naturwissenschaftliches Alltagswissen wird hier angesprochen, also ist die Analyseeinheit der Kategorie „Fachliche Qualifikation“ zuzuordnen, so könnten Koder bei oberflächlicher Betrachtung meinen. Jedem Chinesen, aber auch jedem, der sich mit Geschichte und Landeskunde Chinas intensiv befasst hat, ist bei diesem Text aber klar, dass ebenfalls die Kategorie „Sozialismus als Ideal“ angesprochen wird, denn die kleinen Regentropfen versinnbildlichen die einzelnen Volksgenossen die gemeinsam das große „sozialistische Meer des Volkes“ bilden.

## 5 Theorientest durch empirische vergleichende pädagogische Forschung

Die Datengrundlage dieses Beispiels entstammt der Second International Mathematics Study (SIMS) der International Association for the Evaluation of Educational Achievement (IEA), bei der die Datenerhebung in 18 teilnehmenden Staaten Anfang der achtziger Jahre erfolgte. Ein Ziel internationaler Schulleistungsstudien ist die Feststellung von Schülerleistungen in bestimmten Fächern und die Bestimmung von wesentlichen Einflussfaktoren zum Zwecke des „benchmarkings“ und „systemmonitorings“. Dass diese Daten auch zu weiterführenden Analysen genutzt werden können und auch werden, wird öffentlich und auch fachöffentlich häufig nicht gebührend wahrgenommen.

Im folgenden Beispiel geht es darum, einerseits zu prüfen, ob sich ein in mehreren Untersuchungen bewährter Theorieansatz auch im transkulturellen Vergleich erhärtet, und ob länderspezifische Eigenheiten mittels solcher Analysen deutlich herausgearbeitet werden können. Im Rahmen von Interessens- und Motivationstheoretischen inspirierten Untersuchungen wurde regelmäßig der positive korrelative Zusammenhang zwischen Fähigkeitsselbstkonzept und Leistung, insbesondere der zwischen fachbezogenem Fähigkeitsselbstkonzept und Fachleistung berichtet. Zur transkulturellen Überprüfung dieses Sachverhalts wurden aus dem Schülerhintergrundfragebogen der SIM-Study entsprechende Skalen und Items ausgewählt. Dazu wurde auf ein faktorenanalytisches Verfahren, die Hauptkomponentenanalyse zurückgegriffen, mit der die Eindimensionalität der analysierten Skalen überprüft wird.

Item	Factor 1			Factor 2			Factor 3		
	Parental Support			Self-concept of Mathematical Ability			Importance of Mathematics		
	FRA	JPN	USA	FRA	JPN	USA	FRA	JPN	USA
My Mother thinks that learning maths is very important for me.	.78	.78	.82						
My father thinks that learning maths is very important for me.	.79	.75	.80						
My Parents encourage me to learn as much maths as possible.	.79	.65	.76						
My parents want me to do very well in maths class.	.82	.73	.71						
My parents really want me to do well in mathematics.	.73	.65	.46						
I am not so good at maths.				.79	.75	.80			
Maths is harder for me than for most persons.				.76	.79	.80			
No matter how hard I try I still do not do well in maths.				.81	.68	.82			
I could never be a good mathematician.				.71	.62	.71			
A knowledge of maths is not necessary in most occupations.							.71	.83	.66
Maths is not needed in everyday living.							.77	.82	.72
Most people do not use mathematics in their job.							.69	.73	.61
I can get along well in everyday life without using mathematics							.67	.70	.61

Tabelle 3: Rotierte Faktormatrix von Items zu Einstellungen und Motivation zur Mathematik nach Ländern.

Weisen alle Items der Skala hohe Faktorladungen auf und erklärt der Faktor – die latente gemeinsame Hintergrundvariable der „Einzelitems“ – ein hohes Maß der Antwortvariabilität der Items der entsprechenden Skala, so können die „Einzelitems“ zu einer gemeinsamen Skala zusammengefasst werden, die zu weiterführenden Analysen genutzt werden kann. Aus dem Schülerhintergrundfragebogen konnten so, wie in Tabelle drei ersichtlich, die drei Faktoren „Elterliche Unterstützung“, „Fähigkeitsselbstkonzept Mathematik“ und „Bedeutung von Mathematik“ herausgearbeitet werden. Diese Faktorenlösung ist mit der hier für drei Länder berichteten – Frankreich, Japan und USA – nahezu identisch mit der Faktorenlösung für alle an SIMS teilnehmenden Ländern.

Die Beziehungen dieser Faktoren, erweitert um den Grad der „Ausbildungshöhe des Vaters“ als einem Indikator von „Sozioökonomischem Status“, untereinander und deren Einfluss auf Schülerleistungen in Mathematik wurden nun pfadanalytisch für Frankreich, Japan und die USA untersucht. Die Pfadkoeffizienten sind ähnlich wie andere korrelative Zusammenhänge zu interpretieren. Die Ergebnisse, wie in Abbildung vier ersichtlich, dieses nach statistischen Prüfkriterien für alle drei Länder guten Modells mit erklärter Leistungsvarianz für Frankreich 20%, Japan 11%, USA 26%, bestätigen einer-

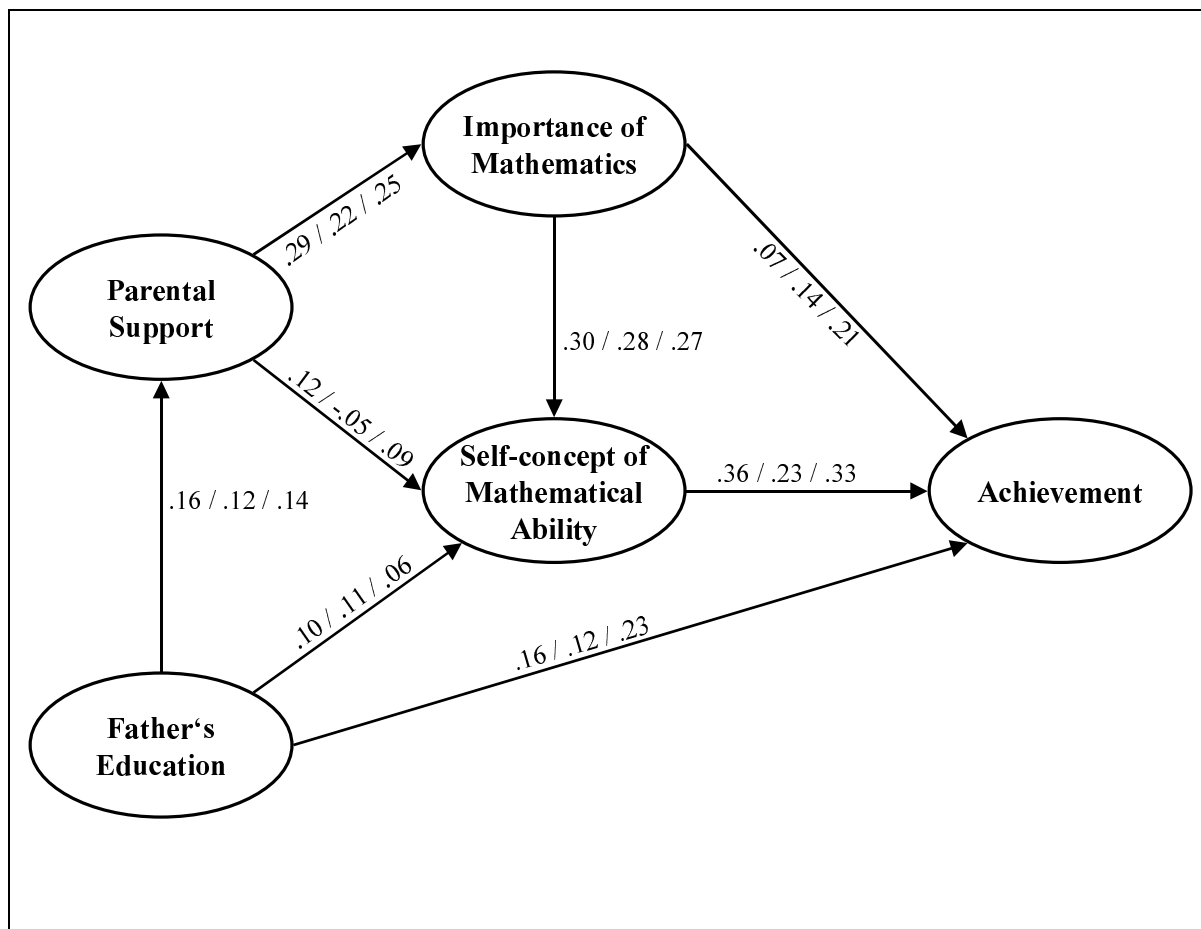


Abbildung 4: Pfadanalyse über die Beziehung motivationaler und familialer Hintergrundvariablen und Fachleistung in drei Ländern (Frankreich/Japan/USA).

die Ergebnisse aus anderen Untersuchungen, nämlich dass das Fähigkeitsselbstkonzept und die Bedeutung, die der Mathematik zugeschrieben wird, starke Moderatorvariablen in Bezug auf die Schülerleistung darstellen. Allerdings sehen wir auch, dass die Beziehungen je nach betrachtetem Land unterschiedlich stark ausgeprägt sind. So ist der direkte Einfluss des Ausbildungsgrades des Vaters in den USA deutlich höher als in Frankreich oder Japan, gleiches gilt für den Einfluss der zugeschriebenen Bedeutung von Mathematik. Bemerkenswert ist aber der negative Zusammenhang von „Elterlicher Unterstützung“ und dem „Fähigkeitsselbstkonzept in Mathematik“ in Japan. Die Autoren dieses Beispiels vermuten hier, dass für ohnehin schon angestrengt arbeitende Schüler, wie den japanischen unterstellt, das „Unterstützungsgebaren“ der Eltern von diesen eher als belastende Anforderung wahrgenommen wird und dass die Eltern sich vielleicht aus diesem Grund mit der offenen Formulierung ihrer „Anforderungen“ zurückhalten.

Mit diesem Beispiel wird vielleicht deutlich, dass im Rahmen vergleichender pädagogischer Forschung theoretische Modelle mittels empirischer Verfahren einerseits auf ihre transkulturelle Validität geprüft, andererseits kulturelle Besonderheiten einzelner Länder deutlich herausgearbeitet werden können. Zur weiterführenden Diskussion dieser nationalen Besonderheiten ist aber wiederum die Berücksichtigung der historischen und

landeskundlichen Dimension notwendig, will die Interpretation dieser Besonderheiten tiefer gehen als hier geschehen.

## 6 Diskussionsbereicherung durch empirische vergleichende pädagogische Forschung

Mit dem letzten kleinen Beispiel soll aufgezeigt werden, wie durch empirisch vergleichende pädagogische Forschung die politische Diskussion angereichert werden kann, sowohl die fachpolitische als auch gesellschaftspolitische. Im Rahmen der Diskussion um interkulturelle Erziehung und um Modelle muttersprachlichen Unterrichts ethnischer Minoritäten gibt es eine Reihe kleinerer Untersuchungen, in denen auch versucht wird, den Sprachstand der verschiedenen Subpopulationen und dessen Bedingungsfaktoren zu erfassen. Dies geschieht häufig mit „entlehnten“ bzw. „selbstgebastelten“ Instrumenten oder gar über Zeugnissenoten, also sehr unzuverlässig. Will man hier solider abgesicherte Ergebnisse erzielen, kann man die international „geeichten“ Instrumente international vergleichender Schulleistungsuntersuchungen nutzen.

Im folgenden wird auf die Ergebnisse einer in Deutschland durchgeführten Zusatzuntersuchung im Rahmen der International Reading Literacy Study der IEA zurückgegriffen, in der Anfang der neunziger Jahre der Sprachstand deutschsprachiger Minderheiten in anderen Ländern mittels eines international validierten und „geeichten“ Testinstruments erfolgte. Im linken Segment der Abbildung fünf findet sich die repräsentative Verteilung der Population der westdeutschen Achtklässler. In der Senkrechten ist die Skala des Leseverständnisses abgebildet, die hier von minus vier bis plus vier reicht. Der Mittelwert der westdeutschen Achtklässler liegt ungefähr bei 1,6, die gesamte westdeutsche Population ist nahezu „normalverteilt“. Auf der unteren Waagerechten ist der jeweilige Prozentsatz der Schüler angegeben, der den auf der Senkrechten ablesbaren Grad an Leseverständnis erreicht hat. Die obere waagerechte Linie markiert eine Grenze, die überschritten werden muss, um ein Leseverständnis erreicht zu haben, das den Mindestanforderungen für eine einfache Berufsausbildung in deutscher Sprache genügt.

Im rechten Segment der Abbildung sind die Vergleichspopulationen getesteter deutscher Minderheiten aus Belgien, Italien (Südtirol), Dänemark, Ungarn und Russland abgebildet. Auf den ersten Blick wird deutlich, dass die getesteten Populationen aus Belgien, Italien und Dänemark einen ähnlich hohen Grad an Leseverständnis erreichen wie die westdeutsche Population. Anders sieht es bei den ungarischen Achtklässlern aus, dort liegt der Mittelwert mit ca. 0,2 deutlich niedriger. Darüber hinaus erreicht rund die Hälfte der Achtklässler der getesteten deutschsprachigen Minderheit in Ungarn nicht den Grad an Leseverständnis, der zu einer einfachen Berufsausbildung in deutscher Sprache vorausgesetzt werden muss. Von der getesteten deutschsprachigen Minderheit in Russland überschreitet nur ein kleiner Anteil diese Grenze der Mindestanforderung.

Wie aus Tabelle vier ersichtlich, werden die Unterschiede im Grad des erreichten Leseverständnisses deutlich, wenn im Rahmen von Kompetenzstufen weiter differenziert wird.

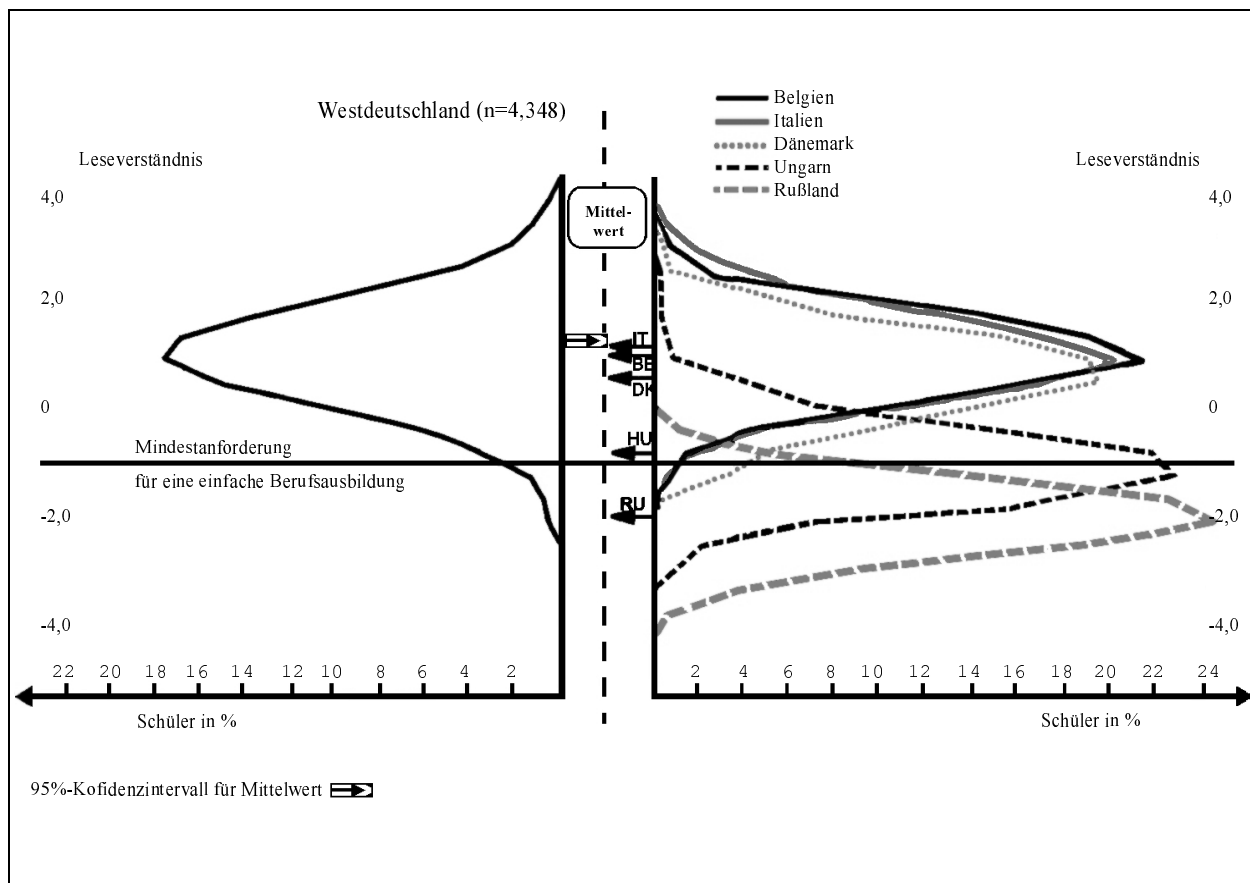


Abbildung 5: Vergleich des Leseverständnisses zwischen Achtklässlern aus Westdeutschland und Minderheitenschulen in Belgien, Italien, Dänemark, Ungarn und Russland.

Stichprobe	Erreichte Stufe des Leseverständnisses		unterhalb der Schwelle minimal funktionalen Leseverständnisses
	evaluativ	generalisiert	
Westdeutschland	63,6%	98,6%	1,4%
Belgien (Deutschsprachige Gemeinschaft)	65,8%	00,0%	0,0%
Italien (Südtirol)	63,3%	99,3%	0,7%
Dänemark (Nordschleswig)	48,1%	98,8%	1,2%
Ungarn (Nationalitätenschulen)	2,1%	51,6%	48,4%
Russland (Rayon Halstadt)	0,0%	3,5%	96,5%

Tabelle 4: Kompetenzstufen des Leseverständnisses bei Achtklässlern aus Westdeutschland und Minderheitenschulen in Belgien, Italien, Dänemark, Ungarn und Russland.

Die deutschsprachigen Minderheiten aus Belgien und Italien erreichen hier ebenfalls den Grad an Leseverständnis, wie die westdeutschen Schüler. Von den getesteten Schülern der deutschsprachigen Minderheit in Dänemark erreicht allerdings ein erheblich kleinerer Anteil als in Westdeutschland, Belgien und Italien die Stufe des „evaluativen Leseverständnisses“.

Für das Verständnis bzw. die Interpretation dieser Befunde bedarf es aber wiederum historischer und landeskundlicher Kenntnisse. So verfügt in Belgien die deutsche Minderheit über ein eigenes Schulwesen, in dem in den ersten zwei Jahren auf „Hochdeutsch“ unterrichtet wird, ab dem dritten Jahr zunehmend in Französisch. In Belgien ist die deutsche Minderheit der flämischen und französischen „Minderheit“ formal gleichgestellt, sie bildet eine eigene politische Gebietskörperschaft mit Dekretgewalt und einem deutschsprachigen Bildungswesen, das vom Kindergarten bis zur Lehrerbildung reicht. Die deutsche Minderheit in Südtirol hat eine relative kulturelle Autonomie mit eigenständigen Minderheitenschulen. In den deutschsprachigen Minderheitenschulen ist Italienisch ab der zweiten bzw. dritten Klasse obligatorisch. Um eine Anstellung im „öffentlichen Dienst“ in Tirol zu bekommen, ist die sehr gute Beherrschung von Deutsch und Italienisch vorgeschrieben, für viele andere Berufe sind gute Deutschkenntnisse notwendig bzw. von Vorteil. In Dänemark organisiert quasi privat der Deutsche Schul- und Sprachverein 25 Kindergärten. Viele dieser Kinder besuchen anschließend eine der 16 deutschen Schulen, in denen sieben Stunden Deutsch und Deutsch auch als Unterrichtssprache unterrichtet wird – es gibt allerdings nur ein Gymnasium. Ab der dritten Klasse ist Reichsdänisch als Fremdsprache verpflichtend. In Dänemark besitzt die deutschsprachige Minderheit kein geschlossenes Siedlungsgebiet, sie konzentriert sich auf Städte und größere Orte. Das zu Hause gesprochene Deutsch ist nicht dem Hochdeutschen zuzurechnen. Deutschkenntnisse haben für eine Berufswahl bzw. -ausübung nur einen geringen Nutzen. In Ungarn gibt es rund 300 deutsche Nationalitätenschulen, in denen ab erstem Schuljahr deutscher Muttersprachenunterricht erteilt wird, allgemeine Unterrichtssprache aber Ungarisch ist. Zwar spielt hier Deutsch bei der Kommunikation im Elternhaus kaum eine Rolle, wohl aber beim Umgang mit deutscher Verwandtschaft und in „deutschorientierten“ Berufen. In Russland (im Altaigebiet) wird an entsprechenden Zusatzschulen zwei- bis sechstündig deutschsprachiger Zusatzunterricht erteilt, in den Elternhäusern wird wenig Deutsch gar „Hochdeutsch“ gesprochen, für eine berufliche Karriere ist Deutsch – von der „Rückwanderung“ einmal abgesehen, irrelevant.

Deutlich wird hier, dass die Testleistung im Leseverständnis in hohem Masse nicht nur vom Modell muttersprachlichen Unterricht, sondern auch von historischen, politischen Rahmenbedingungen und dem Grad der Wertschätzung, der Nützlichkeit abhängt. Diese „Kontextdaten“, hier nur sehr verkürzt erwähnt, sind zwingend bei Interpretationen der Befunde zu berücksichtigen. Mit den Testinstrumenten der IRL-Study stehen hervorragende Instrumente zur Erfassung von Leseverständnis zur Verfügung. Detaillaussagen zu Fragen ethnischer Minoritäten oder gar muttersprachlicher Unterrichtsmodelle für ethnische oder nationale Minderheiten lassen sich mit diesen großen internationalen Schulleistungsuntersuchungen aber nicht diskutieren. Hier ist es umgekehrt wie in den Forschungen, die ich am Anfang dieses Punktes erwähnte. Hier sind die Instrumente gut, aber gezielte Kontextfragen z.B. zur Mehrsprachigkeit von

Schülern und Lehrern oder zu den vielleicht situationsabhängigen Familiensprachen werden nicht gestellt. Vielleicht wäre es hier zielführend, die Testinstrumente der großen Schulleistungsuntersuchungen für Forschungen mit gezielterer Fragestellung zu verbinden, denn so werden die Vorteile beider Untersuchungsarten genutzt ohne die jeweiligen Nachteile in Kauf nehmen zu müssen.

## 7 Schluss

Bevor ich zum Schluss komme, lasse ich es mir natürlich nicht nehmen, einige Worte zur Third International Mathematics and Science Study (TIMSS) zu verlieren. „Deutsche Schülerleistungen sind nur mittelmäßig“, so oder so ähnlich gingen „die“ TIMSS-Ergebnisse durch die Presse. Dieses Benchmarking – immerhin ein realitätsbezogener Vergleich über das, was in schulischen Bildungsprozessen bei vergleichbarer Investition von Zeit und Geld erreicht werden kann – ist ja nur der kleinste, wenn auch vielleicht der populärste Teil der Ergebnisse. Neben zahlreichen Detailanalysen z.B. zur Geschlechtsspezifität der Ergebnisse, zur Abhängigkeit vom sozioökonomischen Status, zur Abhängigkeit von ethnischer Zugehörigkeit etc. liegt eine der Stärken dieser Untersuchung in der Entkräftigung gängiger – und in Deutschland liebgewonnener – bildungspolitisch motivierter Erklärungsmuster. Deutlich wird aus den TIMSS-Befunden, dass die Organisationsstruktur des Bildungswesens nahezu unabhängig von der Schülerleistung in Mathematik und den Naturwissenschaften variiert. Sowohl Gesamtschulsysteme als auch gegliederte Systeme können hervorragende Leistungen erbringen – ich denke hier z.B. an Schweden oder die Schweiz. Sowohl bei zentral als auch bei dezentral organisierter Abschlussprüfung werden Spitzenleistungen erzielt – ich denke hier an Frankreich und die Schweiz. Ob ein Schulsystem Ganz- oder Halbtagschulen bevorzugt, zentral oder dezentral, gegliedert oder integriert organisiert ist, spielt für die Ertragslage nur eine untergeordnete Rolle. Dies zur Kenntnis zu nehmen fällt, insbesondere den Protagonisten der einen oder anderen Variante, schwer. Man kann es deswegen gar nicht laut genug sagen: der deutsche Schulorganisationsstreit ist obsolet. Wir sollten uns lieber mit der wichtigeren Dimension, nämlich mit dem konkreten Unterricht und dessen Verbesserung beschäftigen. Die hier genannten Ergebnisse sind eigentlich nur mittels großer, internationaler Schulleistungsuntersuchungen zu erzielen, da ein einzelnes Land nur selten über die notwendige „Systemvarianz“ verfügt. Allerdings, wie der Unterricht verbessert werden soll, dazu liefern solche Untersuchungen kein Patentrezept.

Ich denke, es ist klar geworden, dass empirische Methoden innerhalb der Vergleichenden Erziehungswissenschaft als erkenntnisgenerierendes Instrument sinnvoll einsetzbar sind, wer sich dem verweigert verschließt die Augen vor der Realität. Dies gilt auf Ebene von Länderstudien ebenso wie beim transnationalen Theorientest mit entsprechenden Auswirkungen auf die vergleichende Theoriediskussion. Klar sollte aber auch geworden sein, dass empirisch vergleichende – gemeint ist hier selbstredend auch die qualitativ empirisch orientierte vergleichende pädagogische Forschung, die hier aus Raumgründen in meinem Vortrag ausgespart wurde – eine landeskundlich und/oder historisch orientierte vergleichende Erziehungswissenschaft nicht ersetzen kann und will. Einander



ergänzend und gezielt kombiniert und mit triangulativem Forschungsdesign können und sollten hier in der vergleichenden pädagogischen Forschung neue Wege beschritten werden.

#### *Quellen:*

Die Beispiele zum Punkt vier, finden sich detailliert in: Bos, W. & Straka, G.A. (1987). Was Maos Erben in der Schule lernen. Ergebnisse einer vergleichenden Inhaltsanalyse von Grundschultextbüchern der VR China. Münster: Waxmann. Das Beispiel zum Punkt fünf ist einem Aufsatz von: Gruehn, S. & Roeder, P.M. (1995). Concomitants of Achievement in Mathematics: A Comparative Analysis. In W. Bos & R.H. Lehmann (Eds.), Reflections on Educational Achievement: Papers in Honour of T. Neville Postlethwaite (pp 88-106). Münster: Waxmann, entnommen. Die Daten und Abbildungen des Beispiels zu Punkt sechs entstammen dem Buch: Lehman, R.H., Peek, R., Pieper, I. & v. Stritzky, R. (1995). Leseverständnis und Lesegewohnheiten deutscher Schüler und Schülerinnen. Weinheim: Beltz. Die Ausführungen unter Punkt sieben zu TIMSS basieren auf: Baumert, J., Bos, W. & Watermann, R. (1998). TIMSS/III. Schülerleistungen in Mathematik und den Naturwissenschaften am Ende der Sekundarstufe II im internationalen Vergleich. Zusammenfassung deskriptiver Ergebnisse. Berlin: Max-Planck-Institut für Bildungsforschung, und: Baumert, J., Bos, W. & Lehmann, R.H. (Eds.). (2000, in Vorbereitung). Dritte Internationale Mathematik- und Naturwissenschaftsstudie. Mathematische und naturwissenschaftliche Bildung am Ende der Schullaufbahn. Band 1: Mathematische und naturwissenschaftliche Grundbildung am Ende der Pflichtschulzeit. Bd. 2: Mathematische und physikalische Kompetenzen am Ende der gymnasialen Oberstufe. Opladen: Leske & Budrich.