



Konstruktion und Evaluation eines Kompetenzstrukturmodells im Bereich mathematischer Repräsentationen

Von der Pädagogischen Hochschule Freiburg
zur Erlangung des Grades
einer Doktorin der Philosophie (Dr. phil.)

genehmigte Dissertation

von

Marianne Bayrhuber-Habeck
aus **Kiel**

Fach: Pädagogische Psychologie

Erstgutachter: Prof. Dr. Markus Wirtz

Zweitgutachter: Prof. Dr. Timo Leuders

Tage der mündlichen Prüfungen: 11. und 17. 11. 2009

Teile dieser Arbeit wurden bereits veröffentlicht unter:

- Bayrhuber, M., Leuders, T., Bruder, R. & Wirtz, M. (angenommen). Repräsentationswechsel beim Umgang mit Funktionen – Identifikation von Kompetenzprofilen auf der Basis eines Kompetenzstrukturmodells. In E. Klieme & D. Leutner (Hrsg.), *Beiheft der Zeitschrift für Pädagogik: Heft 2, 2009. Kompetenzdiagnostik*. Weinheim: Beltz.
- Bayrhuber, M., Leuders, T., Bruder, R. & Wirtz, M. (2009). *Kognitive Strukturen im Bereich Mathematik: Wechsel von sprachlich-symbolhaften Darstellungen und grafisch-bildhaften Repräsentationen*. Vortrag im Rahmen des von Prof. Schnotz und Prof. Baumert organisierten Symposiums "Darstellungsformen im mathematisch-naturwissenschaftlichen Unterricht: Struktur, Entwicklung und Förderung von Repräsentationskompetenz bei Schülern und Lehrern". 72. Tagung der AEPF in Landau, Deutschland, 23.03.-25.03.2009.
- Bayrhuber, M., Leuders, T., Wirtz, M. & Bruder, R. (2009). *Multidimensional or Unidimensional? Assessing and modelling mathematical competence structure*. Presented as individual paper at the EARLI Conference, Amsterdam, The Netherlands, 25.08.-29.08.2009.
- Leuders, T., Wirtz, M., Bayrhuber, M. & Bruder, R. (2009). *A Competence Model for Problem Solving by Using Mathematical Representations*. 33rd Conference of the International Group for the Psychology of Mathematics Education (PME), Thessaloniki, Greece, 19 – 24 July, 2009.
- Bayrhuber, M., Leuders, T., Bruder, R. & Wirtz, M. (2009). *Modellierung und Erfassung mathematischer Kompetenz*. Vortrag am Institut für Psychologie der Universität Heidelberg, Deutschland, 13.01.2009.
- Leuders, T., Bruder, R., Wirtz, M. & Bayrhuber, M. (2008). *DFG-Projekt HEUREKO: Heuristisches Arbeiten mit Repräsentationen funktionaler Zusammenhänge*. Vortrag auf der Herbsttagung des Arbeitskreises ‚Vergleichsuntersuchungen‘ in der Gesellschaft für Didaktik der Mathematik, Freiburg, 23.10. 2008.
- Bruder, R., Wirtz, M., Bayrhuber, M. & Leuders, T. (2008). *Analyse der Kompetenzstruktur im mathematischen Inhaltsbereich „Funktionale Veränderung“*. DFG-Projekt HEUREKO: *Heuristisches Arbeiten mit Repräsentationen funktionaler Zusammenhänge*. Projektvorstellung auf dem Jahreskolloquium des DFG-SPP „Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen“, Darmstadt, Deutschland, 04. – 05. 09. 2008.
- Bayrhuber, M., Leuders, T., Bruder, R. & Wirtz, M. (2008). *Heuristic work with representations of functional coherencies - Diagnosis of the mathematical competencies of students*. Poster presented at the DGF Annual Meeting, Darmstadt, Germany, 04. – 05. September, 2008.

- Bayrhuber, M., Leuders, T., Bruder, R. & Wirtz, M. (2008). *Analyse der Kompetenzstruktur im mathematischen Inhaltsbereich „Funktionale Veränderung“*. Vortrag auf der 71.Tagung der AEPF in Kiel, Deutschland, 25.08.-27.08. 2008
- Wirtz, M., Leuders, T., Bayrhuber, M. & Bruder, R. (2008). *Evaluation of non-unidimensional item contents using diagnostic results from Rasch-analysis*. Paper presented at the EARLI/ Northumbria Assessment Conference (ENAC). Potsdam, Germany, 27-29 August 2008.
- Leuders, T., Bruder, R., Wirtz, M. & Bayrhuber, M. (2008). *Heuristisches Arbeiten mit Repräsentationen funktionaler Zusammenhänge: Diagnose mathematischer Kompetenz von Schülerinnen und Schülern*. Poster präsentiert auf der 71. Tagung der AEPF in Kiel, Deutschland, 25. 08.-27.08. 2008.
- Bayrhuber, M., Leuders, T., Bruder, R. & Wirtz, M. (2008). *Heuristic work with representations of functional coherencies – Assessing and Modelling students competencies*. Poster presented at the Summer Academy on Educational Measurement, IPN, Kiel, Germany, 21.07.-30.07. 2008
- Leuders, T., Bruder, R., Wirtz, M. & Bayrhuber, M. (2008). *Heuristisches Arbeiten mit Repräsentationen funktionaler Zusammenhänge. Diagnose mathematischer Kompetenz von Schülerinnen und Schülern*. Vortrag auf dem Kooperationsstreffen „Aufgabenkonstruktion“ des SPP, Frankfurt, 08. 02. 2008.
- Leuders, T., Bruder, R., Wirtz, M. & Bayrhuber, M. (2007). *Heuristisches Arbeiten mit Repräsentationen funktionaler Zusammenhänge. Diagnose mathematischer Kompetenz von Schülerinnen und Schülern*. Poster präsentiert auf dem Auftakt-Workshop des Schwerpunktprogramms „Kompetenzmodelle“, Duisburg-Essen, 23. – 24.10. 2007.

Inhaltsverzeichnis

1	EINLEITUNG	4
1.1	Kompetenz und Kompetenzmodellierung in der empirischen Bildungsforschung	8
1.1.1	Kompetenzbegriff	8
1.1.2	Kompetenzstrukturmodelle und Kompetenzniveauamodelle	10
1.1.3	Messung von Kompetenzen in nationalen und internationalen „ <i>large scale assessment</i> “-Studien	14
1.2	Mathematikdidaktische Grundlagen	16
1.2.1	Auswahl des Inhaltsbereichs: Wachstum und Veränderung	16
1.2.2	Funktionale Zusammenhänge in der Mathematikdidaktik	18
1.2.3	Repräsentationsformen von Funktionen	20
1.2.4	Curriculare Relevanz verschiedener Darstellungsarten und Repräsentationswechsel bei Funktionen	25
2	STUDIE 1: Entwicklung und Überprüfung eines Kompetenzstrukturmodells für den Umgang mit verschiedenen Repräsentationen von Funktionen im Mathematikunterricht	28
2.1	Theoretischer Hintergrund	28
2.2	Methode	31
2.2.1	Itemkonstruktion	31
2.2.2	Theoretische Modelle	33
2.2.3	Erhebungsdesign	38
2.2.4	Beschreibung der Stichprobe	39
2.2.5	Statistische Verfahren: Ein- und mehrdimensionale Raschmodelle	40
2.3	Ergebnisse	41
2.3.1	Deskriptive Ergebnisse	41
2.3.2	Modellvergleich	44
2.3.3	Überprüfung der Skalenqualität: Reliabilitäts- und Korrelationsanalyse	45
2.4	Diskussion und Ausblick	46

3	STUDIE 2: Modellierung mathematischer Kompetenz anhand von <i>Within-Item-Multidimensionality</i>	49
3.1	Einleitung	49
3.2	Hypothese	51
3.3	Methode	51
3.3.1	Theoretische Modelle	52
3.3.2	Statistisches Verfahren: Within-item-multidimensionality	53
3.4	Ergebnisse	55
3.4.1	Deskriptive Analyse der Lösungswege	55
3.4.2	Zuordnung der Items zu Dimensionen	56
3.4.3	Modellvergleich	58
3.5	Diskussion	59
4	STUDIE 3: Repräsentationswechsel beim Umgang mit Funktionen – Identifikation von Kompetenzprofilen auf der Basis eines Kompetenzstrukturmodells	61
4.1	Theoretischer Hintergrund	61
4.2	Ziele der Studie	64
4.3	Methode	64
4.3.1	Erhebungsdesign	64
4.3.2	Statistisches Analyseverfahren: Latente Klassenanalyse	65
4.3.3	Zusammenhang von Kompetenzprofilen und figuralem Denken	66
4.4	Ergebnisse	67
4.4.1	Analyse von Kompetenzprofilen	67
4.4.2	Zusammenhang von Kompetenzprofilen und kognitiver Leistungsfähigkeit	69
4.5	Diskussion und Ausblick	71
4.6	Grenzen der vorliegenden Studie und Ausblick	72

5	ZUSÄTZLICHE AUSWERTUNGEN	74
5.1	Analyse der Itemschwierigkeiten unter Einbeziehung von Moderatorvariablen anhand von „Differential Item Functioning“	74
5.2	Bestimmung von Kompetenzniveaus in der Dimension „Wechsel von Situation und grafischer Repräsentation“	80
6	GESAMTDISKUSSION	88
6.1	Modellierung mathematischer Kompetenz – ein Fazit	88
6.1.1	Dimensionale Analyse der Kompetenzstruktur	88
6.1.2	Analyse der Kompetenzstruktur anhand von Kompetenzprofilen und Abweichungen von der dimensional Struktur	90
6.2	Implikationen für die Unterrichtspraxis	93
6.3	Schlussfolgerungen und Ausblick	96
7	ZUSAMMENFASSUNG	98
8	LITERATUR	101
9	ANHANG	108
9.1	Itemkennwerte: Pilotierung	109
9.2	Itemkennwerte: Hauptstudie	111
9.3	Skalenkennwerte der Moderatorvariablen (Schülerfragebogen)	113

1 EINLEITUNG

Internationale und nationale Vergleichsuntersuchungen am Ende der Pflichtschulzeit wie z.B. TIMSS (Baumert, Evans & Geiser, 1998) oder PISA (Organisation for Economic Co-operation and Development (OECD), 2005) haben große Beachtung in der Gesellschaft gefunden, weil sie die Diskrepanz zwischen den Zielen des Bildungssystems einerseits und den tatsächlichen Kompetenzen der Schülerinnen und Schüler andererseits sichtbar machen.

Als Konsequenz wurden im Schulwesen weitreichende Maßnahmen zur Kompetenzförderung in Angriff genommen, wie z.B. gezielte Unterrichtsreformen für Mathematik und die Naturwissenschaften (Prenzel, Baumert, Blum, Lehmann, Leutner, Neubrand, Pekrun & Schiefele, 2005) oder Initiativen zur Förderung der Lesekompetenz. Ein Kernelement der neuen Reformagenda ist der Aufbau von Evaluationsagenturen, die Lernergebnisse sowohl auf der Ebene einzelner Klassen und Schulen als auch auf der Ebene von Bildungssystemen ermitteln und zurückmelden.

Für die Optimierung von Bildungsprozessen kommt der Messung von Kompetenzen demnach eine Schlüsselfunktion zu. Thema des DFG-Schwerpunktprogramms „Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen“ ist daher die Erfassung und Modellierung von Kompetenzen im Bildungsbereich. In verschiedenen Domänen werden Kompetenzmodelle und darauf basierende Messinstrumente entwickelt, die es erlauben, Bildungsentscheidungen besser als bisher vorzubereiten. Hierbei handelt es sich z.B. um Entscheidungen hinsichtlich des Zugangs zu einem bestimmten Schultyp oder die Vergabe eines Schulabschlusses und um Förderangebote (vgl. Holling, Vock & Wittmann, 2001).

Die vorliegende Arbeit wurde im Rahmen des Projektes HEUREKO angefertigt, das von der DFG innerhalb dieses Schwerpunktprogramms gefördert wird. Die Zielsetzung ist eine differenzierte Analyse der Kompetenzstruktur von Schüler/innen beim Problemlösen mit verschiedenen Repräsentationen von Funktionen. Es soll ein Kompetenzstrukturmodell entwickelt und Abweichungen systematisch analysiert werden. Zusätzlich sollen Leistungstypen mit bestimmten Leistungsprofilen postuliert werden. Dadurch soll die Grundlage für ein formatives diagnostisches Instrument im Rahmen des Mathematikunterrichts geschaffen werden.

In der fachdidaktischen Literatur spielen mathematische Repräsentationsformen sowie der Wechsel zwischen ihnen eine zentrale Rolle (z.B. Swan, 1985; Ainsworth, Bibby & Wood, 2002; Goldin, 1998). Im Mittelpunkt der Diskussion stehen für den mathematischen Inhaltsbereich „Wachstum und Veränderung“ Tabelle, Graf, Term und Situation (Wortvorschrift) als typische externe Repräsentationsformen von Funktionen (vgl. Malle 2000). Die Interpretation und wechselseitige Übersetzung dieser Darstellungsformen wird als mathematische Schlüsselfähigkeit angesehen und steht daher im Fokus von verschiedenen Lehr- und Lernkonzepten in diesem Bereich (z.B. Barzel, Hußmann & Leuders, 2005; Herget, Malitte & Richter, 2000).

Im mathematischen Inhaltsbereich „Wachstum und Veränderung“ werden in der vorliegenden Arbeit theoriegeleitet Kompetenzmodelle entwickelt und hinsichtlich ihrer Passung zu den empirischen Daten überprüft. Die Analyse erfolgt anhand von Methoden der Item-Response-Theorie (IRT). Zusätzlich soll die Kompetenzstruktur u.a. anhand von Kompetenzprofilen der Schüler/innen aufgeklärt werden. Langfristiges Ziel ist die Bereitstellung eines formativen Diagnose- und Förderinstrumentariums für die Schulpraxis. Das Diagnoseinstrument soll auf der Grundlage des empirisch am besten passenden

Kompetenzstrukturmodells entwickelt werden und kognitive Fähigkeiten von Schüler/innen beim Umgang mit Funktionen strukturiert erfassen. Mögliche Defizite der Lernenden sollen aufgedeckt werden, um gezielt Förderbedarf identifizieren zu können.

Die vorliegende Arbeit ist folgendermaßen aufgebaut:

Nach einer *theoretischen Einleitung*, in der auf die Kompetenzmodellierung in der empirischen Bildungsforschung sowie auf mathematikdidaktische Grundlagen eingegangen wird, werden drei Studien vorgestellt. Die ersten beiden Studien befassen sich mit der dimensionalen Analyse der Kompetenzstruktur.

Die *erste Studie* beinhaltet die Entwicklung und empirische Überprüfung des erwähnten Kompetenzstrukturmodells. Mögliche Kompetenzstrukturmodelle werden aus empirischen Befunden hinsichtlich mathematischer Repräsentationen und Repräsentationswechsel sowie zum Problemlösen abgeleitet. Danach werden die verwendeten Erhebungs- und statistischen Analysemethoden vorgestellt. Schließlich werden die Ergebnisse des Vergleichs verschiedener Kompetenzstrukturmodelle dargestellt und diskutiert.

Die *zweiten Studie* befasst sich mit einer alternativen multidimensionalen Modellierung der Kompetenzstruktur. Dazu werden die Lösungswege der Probanden analysiert und ein neues Kompetenzmodell postuliert, das anhand von IRT-Methoden mit dem am besten passenden Kompetenzstrukturmodell der ersten Studie verglichen wird. Zum Schluss werden die Ergebnisse diskutiert.

Nach der dimensionalen Analyse der Kompetenzstruktur in den ersten beiden Studien, wird die Kompetenzstruktur weiter anhand weiterer Verfahren ausdifferenziert. In der *dritten Studie* werden Kompetenzprofile von Schülerinnen und Schülerin beim Umgang mit

Funktionen auf der Basis des in Studie 1 beschriebenen Kompetenzstrukturmodells identifiziert. Nach einer theoretischen Einleitung werden die Ziele dieser Studie erörtert, die neben einer Analyse von Kompetenzprofilen den Zusammenhang dieser Profile mit kognitiven Variablen umfassen. Im empirischen Teil dieser Studie werden die Analysemethoden beleuchtet, und anschließend werden die Ergebnisse ausgeführt und diskutiert.

Zur weiteren Ausdifferenzierung der Kompetenzstruktur wird in *weiteren Auswertungen* untersucht, ob und inwieweit sich die Schwierigkeiten der Aufgaben in verschiedenen Subgruppen von Schüler/innen unterscheiden. Zusätzlich ist die erste Einteilung einer Dimension des Kompetenzmodells in Kompetenzniveaus dargestellt.

Die Gesamtdiskussion der Ergebnisse zur Modellierung mathematischer Kompetenzstruktur beim Umgang mit Funktionen erfolgt in Kapitel 6. Zusätzlich sind in diesem Kapitel die zentralen Implikationen für die Unterrichtspraxis, die Schlussfolgerungen sowie der Ausblick dargestellt.

1.1 Kompetenz und Kompetenzmodellierung in der empirischen Bildungsforschung

1.1.1 Kompetenzbegriff

Für empirische Studien, die sich mit der Produktivität des Bildungswesens befassen ist der Kompetenzbegriff von zentraler Bedeutung. Insbesondere im Zusammenhang mit der Definition der Ziele von Bildungssystemen und in groß angelegten Erhebungen von Schülerleistungen („large scale assessment“) spielt der Begriff eine bedeutsame Rolle. In den letzten Jahren ist der Kompetenzbegriff zum Gegenstand intensiver Diskussionen in der Psychologie, der Fachdidaktik und der Erziehungswissenschaft geworden (z.B. Csapo, 2004; Weinert, 2001; Klieme, Leutner & Wirth, 2005).

In seiner einflussreichen Expertise zum Kompetenzbegriff beschreibt Franz Weinert (2001) Kompetenzen als „die bei Individuen verfügbaren oder von ihnen erlernbaren kognitiven Fähigkeiten und Fertigkeiten, bestimmte Probleme zu lösen, sowie die damit verbundenen motivationalen, volitionalen und sozialen Bereitschaften und Fähigkeiten, die Problemlösungen in variablen Situationen erfolgreich und verantwortungsvoll nutzen zu können“. Auch wenn Weinert damit eine Verbindung von kognitiven und motivationalen Anteilen formuliert, schlägt er vor, in empirischen Studien, kognitive und motivationale Faktoren getrennt zu erfassen, weil nur so ihre Wechselwirkung erfasst werden kann.

Klieme und Leutner (2006) leiten daraus ihre Ausführungen zum Kompetenzbegriff innerhalb des DFG-Schwerpunktprogramms „Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen“ ab. Kompetenzen werden hier definiert als „kontextspezifische kognitive Leistungsdispositionen, die sich funktional auf Situationen und Anforderungen in bestimmten Domänen beziehen“ (Klieme & Leutner,

2006). Beide Kompetenzbegriffe stehen in direktem Bezug zueinander, sie unterscheiden sich lediglich darin, ob motivationale Anteile eingeschlossen werden (Weinert, 2001) oder ob Kompetenzen auf den kognitiven Bereich beschränkt sind (Klieme & Leutner, 2006). In der vorliegenden Arbeit wird im Sinne einer inhaltlichen Fokussierung der Kompetenzbegriff nach Klieme & Leutner (2006) auf den kognitiven Bereich beschränkt.

Die Kontextabhängigkeit ist das wesentliche Merkmal des hier verwendeten Kompetenzbegriffs. Ein weiteres wichtiges Merkmal des Kompetenzbegriffs nach Klieme & Leutner (2006) ist, dass Kompetenzen durch Erfahrung und Lernen erworben werden und durch Interventionen beeinflusst werden können (vgl. Baumert, Stanat & Demmrich, 2001). Die Kontextabhängigkeit und die Erlernbarkeit von Kompetenzen sind sowohl in der Definition nach Weinert als auch in der Definition nach Klieme und Leutner wesentliche Charakteristika des Kompetenzbegriffs. Anhand dieser Merkmale lässt sich das Konzept der Kompetenz vom klassischen Intelligenzbegriff abgrenzen. Der Kompetenzbegriff wurde als Gegenbegriff zur klassischen Intelligenzforschung in der Psychologie eingeführt, die generalisierte, kontextunabhängige und nur in begrenztem Maße erlernbare kognitive Leistungsdispositionen erforscht (McClelland, 1973).

In der aktuellen Kompetenzforschung in Deutschland herrscht Konsens darüber, dass in empirischen Untersuchungen zur Kompetenz die kognitiven und motivationalen Anteile der Kompetenz getrennt erfasst werden sollten (Ständige Konferenz der Kultusminister der Länder der Bundesrepublik Deutschland (KMK), 2005a). Auch im Schwerpunktprogramm „Kompetenzmodelle“ wird auf die kognitiven Leistungsdispositionen fokussiert (Klieme & Leutner, 2006). In den PISA-Studien werden die motivationalen Aspekte ebenfalls getrennt von den Items zur Messung der drei Bereiche Lesen, Mathematik und Naturwissenschaften anhand von zusätzlichen Fragebogen erfasst (z.B. Baumert et al., 2001).

In der vorliegenden Arbeit wird vor diesem Hintergrund der Kompetenzbegriff von Klieme und Leutner herangezogen. Danach wird die *mathematische Kompetenz im Umgang mit verschiedenen Repräsentationen von Funktionen* als ein Gefüge kognitiver Leistungsdispositionen verstanden, das Schüler/innen dazu befähigt, mit Funktionen anforderungsgemäß umzugehen. Diese Kompetenz wird nicht als von Natur aus gegeben angesehen, sondern muss erlernt werden. Damit ist eine direkte Anknüpfung an das Verständnis von Kompetenz im DFG-SPP „Kompetenzmodelle“ möglich.

1.1.2 Kompetenzstrukturmodelle und Kompetenzniveaumodelle

Es stellt sich bei der empirischen Erfassung von Kompetenzen die Frage, welche Modelle die Grundlage für die Beschreibung von Messergebnissen und für die Entwicklung von Messinstrumenten bilden sollen. Kompetenzmodelle stellen das Bindeglied zwischen dem Konstrukt „Kompetenz“ und den realen Aufgaben dar. In Kompetenzmodellen werden konkrete bereichsspezifische Leistungserwartungen formuliert, wodurch das Verhalten einer Person bei der Bearbeitung von Items eines Tests als Hinweis auf deren Kompetenz zu werten ist (Klieme & Hartig, 2008). Dabei bedürfen Kompetenzmodelle einer empirischen Überprüfung, in der erhoben wird, ob sich die im Modell formulierten Aspekte tatsächlich bei den Schüler/innen nachweisen lassen.

Theoretisch begründete Kompetenzmodelle können dazu dienen einerseits der Konstruktion von Tests, mit deren Hilfe individuelle Stärken und Schwächen der Lernenden aufgedeckt werden können und somit zur Individualdiagnostik und Förderung der Schüler/innen im schulischen Alltag (Klieme, Avenarius, Blum, Döbrich, Gruber, Prenzel, Reiss, Riquarts, Rost, Tenorth & Vollmer, 2003). Andererseits können auf ihrer Basis Kompetenzstrukturen von Gruppen, z.B. Kompetenzprofile von Schüler/innen aufgedeckt werden. Klieme et al. (2003) weist darauf hin, dass es sinnvoller erscheint, einen kleinen Kompetenzbereich detailliert zu erfassen, als ein großes Kompetenzspektrum zu erfassen.

Im Rahmenantrag zur Einrichtung des DFG-Schwerpunktprogramms heißt es: „Die zentrale Aufgabe der Forschung besteht in der Entwicklung von Modellen der Struktur, Stufung und Entwicklung von Kompetenzen, die kognitionspsychologisch fundiert sind und mit avancierten psychometrischen Techniken verbunden werden können. (...) Bei der Formulierung von theoretischen Modellen zur Beschreibung und Erklärung von Kompetenzen geht es zum einen um die Definition von *Strukturen*, zum anderen um die Charakterisierung

von *Niveaus*. *Kompetenzstrukturmodelle* befassen sich mit der Frage, welche und wie viele verschiedene Kompetenzdimensionen in einem spezifischen Bereich differenzierbar sind. Bei der Beschreibung von *Kompetenzniveaus* geht es darum, welche konkreten situativen Anforderungen Personen bei welcher Ausprägung einer Kompetenz bewältigen können“ (Klieme & Leutner, 2006).

Bei *Kompetenzstrukturmodellen* steht die Frage nach der Dimensionalität von Kompetenzstrukturen im Mittelpunkt. Die Zusammenhänge von Kompetenzen in verschiedenen Bereichen kann Gegenstand dieser Modelle sein. Es kann eingeschätzt werden, wie die Bewältigung unterschiedlicher Anforderungen miteinander korreliert und auf welchen Dimensionen Kompetenzunterschiede beschrieben werden können. Die Entscheidung für ein bestimmtes Strukturmodell, d.h. wie differenziert einzelne Kompetenzbereiche betrachtet werden, sollte von theoretischen und ökonomischen Überlegungen begleitet werden. Auch die differenzierte Diagnostik von Teilkompetenzen kann anhand von Strukturmodellen erfolgen. In diesem Fall wird die Binnenstruktur einzelner Kompetenzbereiche, z.B. mathematische Fähigkeiten, hinsichtlich zugrunde liegender Teilkompetenzen und ihrer Zusammenhängen analysiert. In *Kompetenzniveaumodellen* werden spezifische Fähigkeiten bei unterschiedlichen Ausprägungen einer Kompetenz beschrieben (vgl. Hartig, 2007).

Ein typisches Kompetenzstrukturmodell strukturiert eine Kompetenz in Kompetenzdimensionen und unterteilt jede dieser Dimensionen in Kompetenzniveaus. Die Dimensionen eines Kompetenzmodells repräsentieren inhaltlich unterscheidbare Aspekte anhand derer sich Kompetenzunterschiede zwischen Personen beschreiben lassen. Hierbei handelt es sich sowohl um inter- als auch (bei der Abbildung von Kompetenzentwicklungen) um intraindividuelle Unterschiede (Klieme, Hartig, Kenk & Fleischer, 2009).

In der vorliegenden Arbeit wird ein Kompetenzstrukturmodell hinsichtlich des Umgangs mit verschiedenen Repräsentationen von Funktionen abgeleitet und empirisch überprüft. Dadurch wird die Grundlage geschaffen, Unterschiede von Schülerinnen und Schülern beim Problemlösen mit Funktionen zu erfassen und so aufzudecken, an welcher Stelle eine differenzierte Förderung anzusetzen hat. Zur Überprüfung des Modells kommen u.a. mehrdimensionale Rasch-Modelle (Wu & Adams, 2006) und ein Signifikanztest zum Einsatz. Zusätzlich werden erste Analysen der Kompetenzstruktur abgeleitet und Kompetenzprofile von Schüler/innen beim Umgang mit Funktionen untersucht.

1.1.3 Messung von Kompetenzen in nationalen und internationalen „large scale assessment“-Studien

In der empirischen Bildungsforschung wurden in den letzten Jahren – unter Beteiligung deutscher Wissenschaftler – komplexe psychometrische Modelle entwickelt, die es erlauben, Item- und Personenmerkmale, kategoriale Unterscheidungen und Kompetenzentwicklungen abzubilden (z.B. Wilson, De Boeck & Carstensen, 2008; Rupp, 2005).

In der US-amerikanischen Bildungsforschung finden in jüngster Zeit zunehmend IRT-Modelle mit kategorialen latenten Variablen Verwendung (z.B. von Davier, 2005; Wilson, 2003). In diesen Modellen werden interindividuelle Unterschiede anhand von qualitativen Kategorien abgebildet (z.B. Latente Klassenanalyse oder Mischverteilungsmodelle, z.B. Rost, 2003; Muthen & Muthen, 2000). Wenn Unterschiede zwischen Gruppen von Schüler/innen anhand von quantitativen Dimensionen erfasst werden, kommen IRT-Modelle mit kontinuierlichen Variablen zum Einsatz (z.B. Raschmodell, LLTM, vgl. Rost & Carstensen, 2002) oder auch mehrdimensionale Raschmodelle (Adams, Wilson & Wang, 1997).

In mehrdimensionalen Rasch-Modellen wird zur Entscheidung, wie viele Dimensionen unterscheidbar sind, ein Vergleich der Modellgütekriterien des mehrdimensionalen Modells mit denen des eindimensionalen Modells herangezogen. Zusätzlich werden die Korrelationen zwischen den Dimensionen beurteilt (z.B. Hartig, 2007). Die einzelnen Dimensionen werden meist in Kompetenzniveaus unterteilt und dienen einer qualitativen Beschreibung der Anforderungen, die Schüler/innen mit unterschiedlich stark ausgeprägten (Teil-) Kompetenzen bewältigen können (vgl. Kapitel 1.1.2). Dieses Prinzip, Leistungen in Kompetenzniveaus einzuteilen, wurde erstmals in den USA im Rahmen des National Assessment of Educational Progress (NAEP) entwickelt. Die Einteilung erfolgte

post hoc auf der Basis der empirisch bestimmten Itemschwierigkeiten und anhand von eher intuitiver Charakterisierungen der Niveaus („*scale anchoring*“, Beaton & Allen, 1992).

Dieses Verfahren wurde – leicht abgewandelt- in TIMSS und auch in PISA verwendet.

Im Rahmen der DESI-Studie entwickelte Hartig (2007) theoriegeleitet eine empirische Abgrenzung von Niveaustufen, wobei die Anzahl und die Breite der Kompetenzniveaus nach empirischen Kriterien bestimmt werden (vgl. Kapitel 5.2). Wilson (2005) verwendet in seinem Ansatz zur theoriegeleiteten Entwicklung kontinuierlicher Dimensionen sogenannte „Construct Maps“. In der neueren psychometrischen Forschung kommen Ansätze zur Anwendung, die Aufgaben- und Personenmerkmale auf der Basis des Linear-Logistischen Testmodells (LLTM) kombinieren, um Schülerleistungen in Kompetenzniveaus einzuteilen (vgl. Wilson, de Boeck & Carstensen, 2008). In nationalen und internationalen „*large scale assessment*“-Studien werden Kompetenzen also zumeist anhand von IRT-Modellen erfasst und empirisch zugänglich gemacht. Auf ihrer Basis können für die Schulpraxis relevante Rückmeldungen über die Ausprägung und die Veränderung von Kompetenzen gegeben werden.

1.2 Mathematikdidaktische Grundlagen

1.2.1 Auswahl des Inhaltsbereichs: Wachstum und Veränderung

Zur Ableitung und empirischen Überprüfung eines Kompetenzstrukturmodells hinsichtlich des Umgangs mit verschiedenen Repräsentationen von Funktionen wird zunächst der Inhaltsbereich eingegrenzt. Es wird der Fachinhaltsbereich „mathematische Beschreibung von Wachstum und Veränderung“ gewählt, da dieser von besonderer allgemeinbildender Bedeutung ist und national und international im Mathematikunterricht stark repräsentiert ist. Zusätzlich zeichnet er sich durch theoretische und z.T. auch empirische Konsolidiertheit aus.

So umfasst der gewählte Inhaltsbereich ein zentrales Thema der Mathematik: die der funktionalen Abhängigkeit. Die besondere Rolle „funktionalen Denkens“ ist in der Mathematikdidaktik bereits seit über hundert Jahren präsent (vgl. Klein & Schimmack, 1907; Krüger, 1999). Zusätzlich lassen sich im Bereich „Wachstum und Veränderung“ die von Winter (1995) geforderten allgemeinbildenden Grunderfahrungen im Mathematikunterricht realisieren: (i) zur Erfahrung der Anwendung von Mathematik (Modellieren), (ii) zur Erfahrung innermathematischer Strukturierungsleistung, (iii) und zur Erfahrung aktiv-problemlösenden Arbeitens. Der so beschriebene allgemeinbildende Charakter des Inhaltsbereiches spiegelt sich außerdem in der zentralen Rolle wieder, die ihm national und international in normativen Kompetenzmodellen zugewiesen wird (Niss, 2003).

Es wird die Existenz typologischer Strukturen oder Kompetenzprofile von Schüler/innen in dem gewählten Inhaltsbereich aufgrund fachdidaktischer Überlegungen vermutet. Daher ist die Verletzung einfacher eindimensionaler Modelle anzunehmen. Für ein diagnostisches Instrument, das den gewählten Inhaltsbereich erfassen soll, ist die Kenntnis einer validen dimensional Struktur erforderlich.

Nach der Wahl und Begründung des Inhaltsbereichs „Wachstum und Veränderung“ wird im nächsten Abschnitt auf funktionale Zusammenhänge eingegangen, die zentral für diesen Fachinhaltsbereich sind.

1.2.2 Funktionale Zusammenhänge in der Mathematikdidaktik

Dem Einsatz von Funktionen kommt im Mathematikunterricht eine vielfältige Bedeutung zu. Zum einen können mit Funktionen abstrakte Zusammenhänge visualisiert werden, andererseits dienen sie der Vermittlung mathematischer Denk- und Arbeitsweise im Unterricht. Entsprechend der Bedeutung von funktionalem Denken im Mathematikunterricht ist die Vermittlung dieser Kompetenzen in den deutschen Lehrplänen und in den Bildungsstandards für Mathematik (KMK, 2005) verankert.

Der Funktionsbegriff wird in der Mathematikdidaktik seit Anfang des letzten Jahrhunderts diskutiert (vgl. Lietzmann 1909 zur Meraner Reform, zit. nach Krüger, 1999), modernere Ansätze gehen auf die Arbeiten von Vollrath (1989), Malle (2000) sowie Leuders und Prediger (2005) zurück. Wesentlich für den Umgang mit Funktionen sind drei grundlegende Aspekte (Vollrath, 1989):

- *Der Zuordnungsaspekt:* „Durch Funktionen beschreibt oder stiftet man Zusammenhänge zwischen Größen: einer Größe ist dann eine andere zugeordnet, so dass die eine Größe als abhängig gesehen wird von der anderen“.
- *Das Änderungsverhalten:* „Durch Funktionen erfasst man, wie Änderungen einer Größe sich auf eine abhängige Größe auswirken. [...] Die Ausprägung des funktionalen Denkens zeigt sich also auch daran, in welcher Weise Änderungen geplant, durchgeführt, analysiert und zur Lösung von Problemen eingesetzt werden können.“
- *Die Sicht auf die „Funktion als Ganzes“:* „Mit Funktionen betrachtet man einen gegebenen oder erzeugten Zusammenhang als Ganzes. [...] Die Ausprägung des funktionalen Denkens zeigt sich an der Fähigkeit, in unterschiedlichen Darstellungen von Funktionen das Ganze der Funktion zu erfassen und in der Fähigkeit, vom Einzelnen aufs Ganze und umgekehrt vom Ganzen aufs Einzelne ‚umzuschalten‘“.

Mit dem sogenannten Zuordnungsaspekt fasst Vollrath die statischen Elemente funktionalen Denkens zusammen, während das *Änderungsverhalten* die dynamischen Aspekte beinhaltet. Dieser Änderungsaspekt wird auch als Kovariation bezeichnet (Malle, 2000). Das ganzheitliche Denken wird nach Vollrath mit dem dritten Aspekt, der Sicht auf die *Funktion als Ganzes*, d.h. die Menge aller Wertepaare oder den Grafen als Gestalt, erfasst. Die Funktion wird als eigenständiges Objekt betrachtet, ein Beispiel wäre der Sinus als charakteristischer periodischer Graf einer harmonischen Schwingung oder die symbolische Darstellung als Term (vgl. Leuders & Prediger, 2005).

Diese Aspekte von Funktionen sind für die Charakterisierung von Items des entwickelten Tests von Bedeutung und werden in Abschnitt 5.2 bei der Einteilung einer Dimension des Kompetenzmodells berücksichtigt.

1.2.3 Repräsentationsformen von Funktionen

Funktionen können auf unterschiedliche Weise extern repräsentiert sein (vgl. Malle 2000). Neben der Darstellung in einer Tabelle, die bereits im Mathematikunterricht in der Grundschule erlernt wird, können funktionale Zusammenhänge auch grafisch, verbal oder als Funktionsterm dargestellt sein. Jede dieser Darstellungsformen weist mehr oder weniger deutlich Aspekte funktionalen Denkens nach Vollrath (1989) auf. Zur Illustration der verschiedenen Darstellungsarten sind (außer bei der symbolischen Repräsentation) jeweils Beispiele aus dem in der vorliegenden Arbeit entwickelten Mathematiktest aufgeführt.

Tabellarische Darstellung

Ein funktionaler Zusammenhang kann als Tabelle dargestellt sein, wobei der Zuordnungsaspekt besonders gut erkennbar ist, wenn die Tabelle spaltenweise gelesen wird. Es kann zu jedem x der dazugehörige Funktionswert y ermittelt werden. Den Änderungsaspekt können die Schülerinnen und Schüler erkennen, wenn sie die Tabelle zeilenweise lesen. Der Objektaspekt ist für die Lernenden in der tabellarischen Darstellung schwer erkennbar, da nur ein Ausschnitt eines funktionalen Zusammenhangs dargestellt ist. Das folgende Beispiel in Abbildung 1 illustriert die tabellarische Repräsentationsform einer Funktion.

Abb.1: Tabellarische Darstellung

Eine Bakterienkultur bedeckt am Anfang (also am Tag 0) $0,5 \text{ cm}^2$ eines Nährbodens.
 Sie vermehrt sich so schnell, dass sich die bedeckte Fläche von Tag zu Tag verdoppelt.



Ergänze die Tabelle!

<i>Anzahl der Tage</i>					
<i>Größe der bedeckten Fläche in cm^2</i>					

Algebraische Darstellung

Mit der algebraischen Darstellung eines funktionalen Zusammenhangs ist eine Darstellung als Term unter Verwendung von Variablen als algebraische Symbole gemeint. Damit die Lernenden die drei Aspekte Vollraths (1989) erkennen können, ist ein solides Wissen über das Variablenkonzept erforderlich, welches sehr anspruchsvoll ist (vgl. Malle, 2000). Dieses wird meist erst nach längerer Beschäftigung mit Funktionen entwickelt, ein Beispiel findet sich in der folgenden Abbildung.

Abb.2: Algebraische Darstellung

Mit der Formel

$$y = (x - 32) \cdot 5 : 9$$

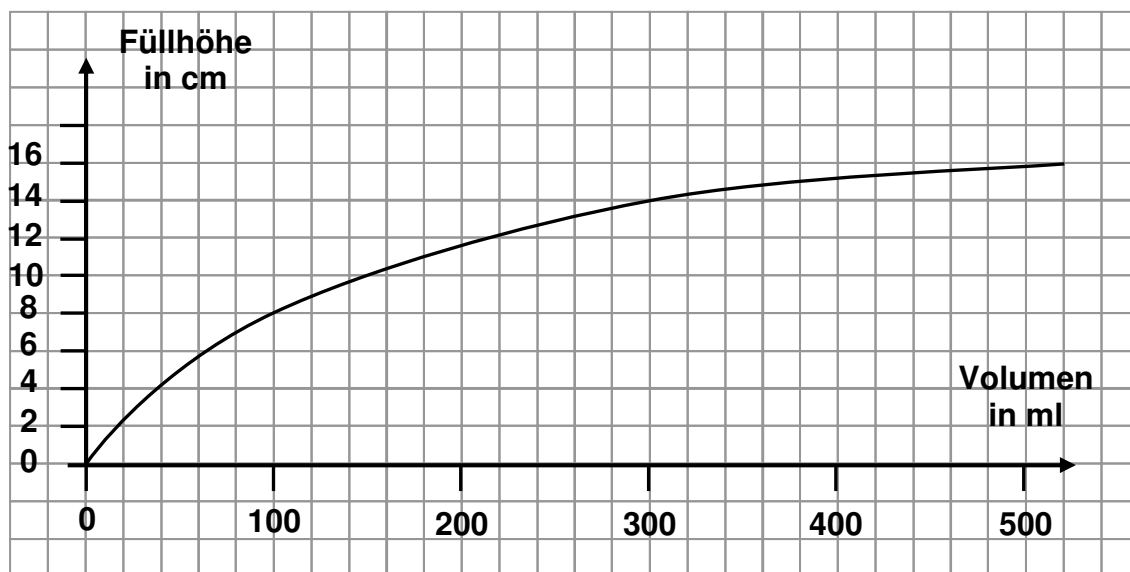
kann man amerikanische Temperaturangaben, die in Fahrenheit angegeben werden, in die in Europa gebräuchliche Celsiusangabe umrechnen. Erstelle eine Anleitung für Personen, die nicht mit Termen umgehen können.

Grafische Darstellung

In der grafischen Darstellung einer Funktion ist der Zuordnungsaspekt gut erkennbar, wenn für ein bestimmtes x der zugehörige y -Wert abgelesen wird. Der Änderungs- bzw. der Kovariationsaspekt (Malle, 2000) lässt sich feststellen, wenn die Schüler/innen erkennen, dass sich die Funktion von x in bestimmter Weise ändert, wenn sich x verändert. Nach Vollrath (1989) ist der Objektaspekt insbesondere bei der grafischen Darstellung zu erkennen, da Eigenschaften von Funktionen wie Monotonie oder Symmetrie hier besonders auffallen. In Abbildung 3 ist eine grafische Darstellung einer Funktion gezeigt. Die besondere Bedeutung der grafischen Darstellung wird in einigen didaktischen Konzepten hervorgehoben (z.B. Swan, 1985; Herget et al., 2000).

Abb.3: Grafische Darstellung

Für einen Haushaltsmessbecher ist der Zusammenhang zwischen Füllhöhe und Volumen im Diagramm dargestellt.



Wie hoch stehen 360 ml Wasser?

Antwort: Das Wasser steht _____ cm hoch.

Situative Darstellung

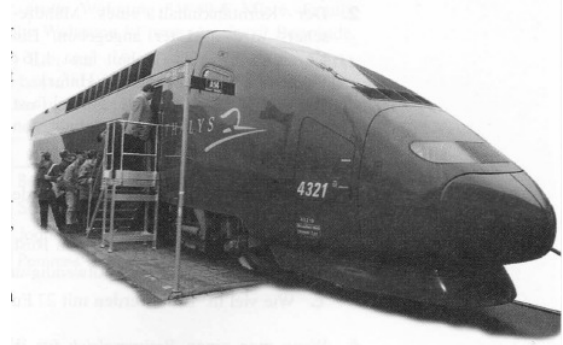
Unter einer situativen Darstellung wird die verbale oder ikonische Darstellung einer Realsituation, die noch keine mathematische Symbole oder Strukturen verwendet, verstanden.

Die drei Aspekte Vollraths (1989) sind in dieser Repräsentationsform besonders schwer zu erkennen, da der funktionale Zusammenhang noch nicht mathematisiert vorliegt. In

Abbildung 4 findet sich die situative Darstellung einer Funktion.

Abb.4: Situative Darstellung

Zwei Testzüge werden auf einer Versuchsstrecke getestet. Sie fahren jeweils mit etwa gleichbleibender Geschwindigkeit ohne Zwischenhalt. Testzug Nr. 1 benötigt für 40 km eine Fahrzeit von 15 min. Testzug Nr. 2 legt 45 km in 25 min zurück.



Welcher Zug fährt schneller?

Löse mithilfe von Tabellen.

Ergebnis: Testzug Nr. _____ ist schneller.

Die Darstellungsarten von Funktionen und der Wechsel zwischen unterschiedlichen Repräsentationsformen wurden in verschiedenen empirischen Studien untersucht. Die Mehrheit dieser Studien, sind Interventionsstudien, welche die Wirksamkeit konkreter Lehr-Lernmodelle untersuchen. Eine theoretisch gut begründete und in verschiedenen empirischen Untersuchungen bestätigte Aussage besteht darin, dass verschiedene Darstellungsformen mathematischer Zusammenhänge und deren Wechsel den individuellen Lernprozess fördern können (Seufert, 2003; Ainsworth et al., 2002). Eine Voraussetzung für den Fördereffekt ist jedoch, dass die mit den Darstellungsformen verbundenen Vorstellungen weiterführendes Lernen nicht behindern (vgl. Gagatsis & Elia, 2004). Zusätzlich müssen die Schülerinnen und Schüler in der Lage sein, zwischen den vorliegenden Repräsentationen zu wechseln, die verbindenden Elemente zu identifizieren und zu verknüpfen (Seufert, 2003; Ainsworth et al., 2002).

1.2.4 Curriculare Relevanz verschiedener Darstellungsarten und Repräsentationswechsel von Funktionen

Der im letzten Abschnitt verdeutlichten besonderen Funktion mathematischer Darstellungsformen für die Entwicklung mathematischer Allgemeinbildung entspricht auch eine umfassende Berücksichtigung dieses Kompetenzaspektes in den Bildungsstandards für das Fach Mathematik (z.B. KMK, 2005). Am Ende der zehnten Klasse sollen Schülerinnen und Schüler Kompetenzen erworben haben, die sie dazu befähigen, Funktionen flexibel einzusetzen, zu analysieren und zu konstruieren:

- **(K 4) Mathematische Darstellungen verwenden**

[Die Schülerinnen und Schüler können] „verschiedene Formen der Darstellung von mathematischen Objekten und Situationen anwenden, interpretieren und unterscheiden; Beziehungen zwischen Darstellungsformen erkennen; unterschiedliche Darstellungsformen je nach Situation und Zweck auswählen und zwischen ihnen wechseln“.

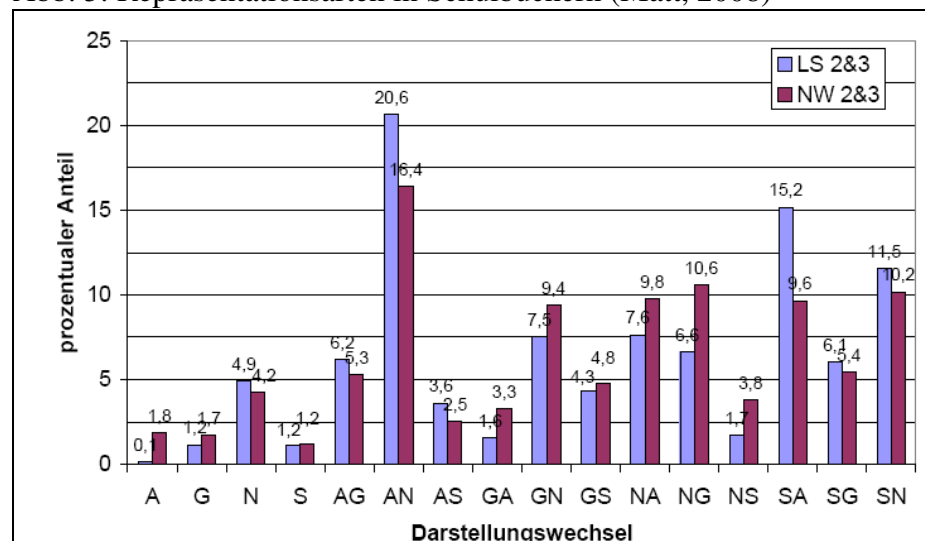
- **(L 4) Leitidee Funktionaler Zusammenhang**

„Die Schülerinnen und Schüler nutzen Funktionen als Mittel zur Beschreibung quantitativer Zusammenhänge; erkennen und beschreiben funktionale Zusammenhänge und stellen diese in sprachlicher, tabellarischer oder grafischer Form sowie gegebenenfalls als Term dar; analysieren, interpretieren und vergleichen unterschiedliche Darstellungen funktionaler Zusammenhänge (wie lineare, proportionale und antiproportionale).“

In Schulbüchern für den Mathematikunterricht (z.B. Böttner, Maroska, & Olpp, 2005) und Lehrplänen (vgl. KMK, Hessen. Lehrplan Mathematik, 2005; Ministerium für Kultus, Jugend und Sport Baden-Württemberg: Bildungsplan für das Gymnasium der Normalform, 2004) werden die verschiedenen mathematischen Darstellungen von Funktionen und der Wechsel der Repräsentationsart gemäß den Bildungsstandards umfassend berücksichtigt.

Diese curriculare Bedeutung schlägt sich auch in der Struktur gängiger Schulbücher wieder. In Abbildung 5 sind die prozentualen Anteile der Aufgaben in einer bestimmten Repräsentationsform in den Schulbüchern der Reihe „Lambacher Schweizer“ (LS) und der Reihe „Neue Wege“ (NW) dargestellt (Matt, 2008). Es sind die folgenden Darstellungsarten von Funktionen aufgeführt: Algebraische Darstellung (A), grafische Darstellung (G), numerische Darstellung (N) die situative Darstellung (S), sowie der Wechsel zwischen diesen Repräsentationsraten (z.B. AN, GA).

Abb. 5: Repräsentationsarten in Schulbüchern (Matt, 2008)



Die Analyse der Schulbuchaufgaben in den Reihen „Lambacher Schweizer“ (LS) und „Neue Wege“ (NW) hat ergeben, dass beide Reihen den Schwerpunkt auf den Wechsel zwischen der algebraischen und numerischen Repräsentation legen. Es scheint ein Konsens

darüber zu bestehen, dass dem Wechsel zwischen der algebraischen und numerischen Darstellung (AN und NA) die größte Bedeutung zukommt. Auch der Wechsel zwischen der situativen und grafischen Darstellung (GS und SG) sowie der Wechsel zwischen der situativen Darstellung und der numerischen Repräsentation (SN und NS) scheinen wesentlich, jedoch nicht in dem Maße wie der Wechsel zwischen der algebraischen und numerischen Darstellung. Es ist fraglich, ob diese Gewichtung aus fachdidaktischer Sicht sinnvoll ist und ob nicht eine höhere Ausgeglichenheit der Wechsel zwischen den verschiedenen Darstellungsformen von Funktionen anzustreben wäre.

2 STUDIE 1:

Entwicklung und Überprüfung eines Kompetenzstrukturmodells für den Umgang mit verschiedenen Repräsentationen von Funktionen im Mathematikunterricht

2.1 Theoretischer Hintergrund

Nationale und internationale *large scale assessment*-Studien wie TIMSS oder PISA (OECD, 2005, Baumert et al., 2001) verwenden meist das eindimensionale Raschmodell, da angenommen wird, dass die Testresultate der Schüler/innen anhand einer eindimensionalen Fähigkeitsdimension beschrieben werden können. Diese eindimensionalen Modelle sind nützlich, etwa wenn das Leistungsniveau von großen Gruppen in verschiedenen Domänen untersucht werden soll. Trotzdem weisen diese Modelle einige Grenzen auf, wenn die Leistungen in komplexeren Kompetenzbereichen untersucht werden soll oder kognitive Prozesse modelliert werden sollen. In der vorliegenden Studie soll mathematische Kompetenz beim Problemlösen mit verschiedenen Repräsentationen von Funktionen modelliert werden. Folgende Fragestellungen sollen untersucht werden:

Sind der Repräsentationstyp und der Wechsel von Repräsentationen wesentlich für das Problemlösen mit Funktionen? Können eine oder mehrere Dimensionen beim Problemlösen mit Funktionen identifiziert werden?

Mathematisches Problemlösen lässt sich definieren als „the cognitive process of figuring out how to solve a mathematical problem that one does not already know how to solve“ (Mayer & Hegarty, 1996, S.31). Beim Problemlösen ist es notwendig, verschiedene kognitive Prozesse zu integrieren. Die Schüler/innen müssen zunächst die verbale Information

einer Aufgabe verstehen, dann müssen sie eine angemessene mentale Repräsentation der Aufgabe bilden sowie einen Lösungsweg planen und überwachen. Schließlich muss die Lösung angemessen prozedural ausgeführt werden. (Desoete, Roeyers & De Clercq, 2003).

Beim Problemlösen mit Funktionen im Mathematikunterricht spielen die verschiedenen Repräsentationsformen von funktionalen Zusammenhängen eine Rolle (vgl. Swan 1985; Ainsworth et al., 2002; Goldin, 1998). Im Mittelpunkt der Diskussion stehen für den mathematischen Inhaltsbereich „Wachstum und Veränderung“ Tabelle, Graf, Term und Situation (Wortvorschrift) als typische externe Repräsentationsformen von Funktionen (vgl. Malle 2000). Die Interpretation und wechselseitige Übersetzung dieser Darstellungsformen wird als mathematische Schlüsselfähigkeit angesehen und steht daher im Fokus von verschiedenen Lehr- und Lernkonzepten in diesem Bereich (z.B. Barzel et al., 2005; Herget et al., 2000). Trotz der hohen Relevanz innerhalb der Mathematikdidaktik, steht eine systematische empirische Untersuchung des Wechsels zwischen „Tabelle“ bzw. „Term“ und anderen externen Repräsentationsformen von Funktionen noch weitgehend aus.

Es hat sich in empirischen Untersuchungen zum Multi-Media-Lernen gezeigt, dass unterschiedliche Repräsentationsformen mathematischer Zusammenhänge und deren Wechsel entscheidend für individuelle Lernprozesse sind. Damit sich ein positiver Lerneffekt einstellt, müssen die Schülerinnen und Schüler in der Lage sein, zwischen den vorliegenden Repräsentationen zu wechseln, die verbindenden Elemente zu identifizieren und zu verknüpfen (Seufert, 2003; Ainsworth et al., 2002). Allerdings zeigen sich hier immer wieder Schwierigkeiten (vgl. Bodemer, Ploetzner, Feuerlein & Spada, 2004). Kozma und Russell (1997) berichten, dass Novizen verschiedene Repräsentationen nur über deren Oberflächenmerkmale aufeinander beziehen, Experten hingegen stellen Bezüge über dahinterliegende Konzepte her.

Wie diese kognitiven Prozesse ablaufen können, d.h. wie Informationen, die aus Texten und Grafen entnommen wurden, zueinander in Beziehung gesetzt werden, wird im Modell zur integrativen Bild- und Textverarbeitung (Schnotz, 2005) postuliert. Es werden die folgenden Verarbeitungsstufen der Bild-Text-Integration unterschieden: (a) Zunächst werden Detailinformationen aus Bild und Text miteinander verknüpft. (b) Danach werden einfache semantische Relationen zwischen Text und Bild erkannt und auf die jeweils andere Repräsentationsform bezogen. (c) Schließlich werden komplexe Relationsgefüge von Text und Bild verknüpft.

Ausgehend von dieser Literaturlage erscheint es plausibel, eine differenzierte Kompetenzstruktur im Umgang mit mathematischen Repräsentationsformen anzunehmen und diese in Form eines Kompetenzstrukturmodells empirisch zu testen. Im curricular zentralen Bereich der Leitidee „Wachstum und Veränderung“ werden in der vorliegenden Studie mögliche Kompetenzstrukturmodelle abgeleitet und bezüglich ihrer Passung zu den empirischen Daten überprüft. Auf Grundlage des empirisch am besten passenden Kompetenzstrukturmodells soll ein Diagnoseinstrument entwickelt werden, das kognitive Fähigkeiten von Schülern beim Umgang mit Funktionen strukturiert erfasst. Mögliche Defizite sollen aufgedeckt werden, um gezielt Förderbedarf identifizieren zu können.

2.2 Methode

2.2.1 Itemkonstruktion

Die Items wurden auf der Basis der Schulbuchaufgaben entwickelt, um eine möglichst hohe curriculare Validität zu gewährleisten. Anhand einer Aufgabenanalyse in gängigen Lehrbüchern der 7. und 8. Klasse aus Realschulen und Gymnasien konnten vier theoretisch relevante Repräsentationsformen von Funktionen identifiziert werden: Der Wechsel von der grafischen und der situativen Darstellung (Wortvorschrift oder bildliche Darstellung einer Situation) sowie der Wechsel von der numerischen und der situativen Darstellung und die Verarbeitung innerhalb der numerischen, bzw. grafischen Darstellung. Ein Beispiel für den Wechsel zwischen der situativen und grafischen Darstellung findet sich in der folgenden Abbildung 6.

Abb. 6: Aufgabenbeispiel

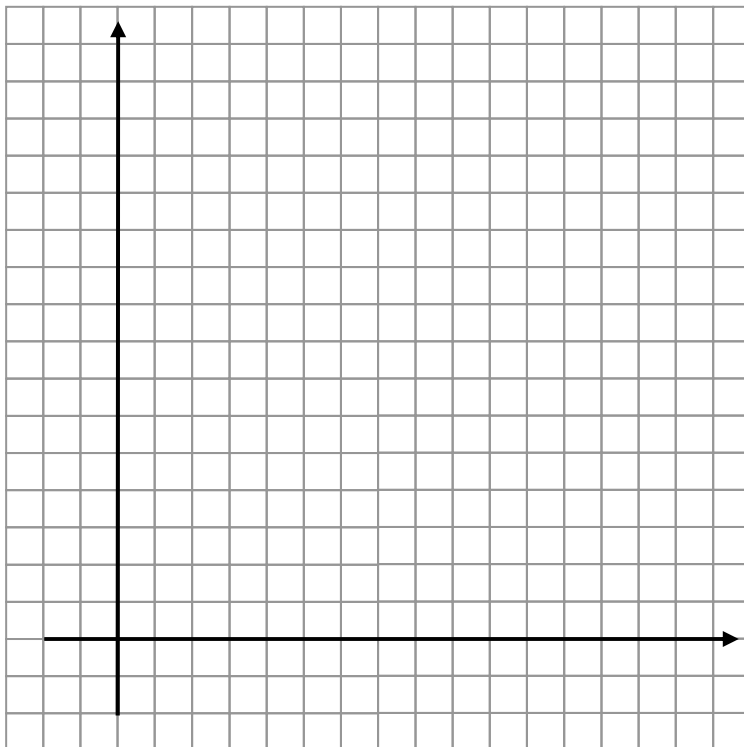


Sonderangebot

**Extrem lange
Brenndauer!**

**15 cm
9 Stunden**

Wie lang ist die Kerze nach 6 Stunden Brennzeit?
Löse mithilfe eines Schaubildes.



Antwort: Nach 6 Stunden ist die Kerze _____ cm lang.

Zusätzlich zu den als wesentlich erachteten Repräsentationsformen wurden die Handlungstypen „Erkennen“, „Beschreiben“ und „Begründen“ für die Itemkonstruktion verwendet, um wesentliche Schülerhandlungen im Untersuchungsbereich zu charakterisieren.

Die beiden Aspekte „Repräsentationstyp“ und „Handlungstyp“ wurden in einer zweidimensionalen Matrix zusammengefasst, die als Basis zur Aufgabenkonstruktion diente. Es wurden insgesamt 80 Items konstruiert, die jeweils einer Zelle dieser Matrix zugeordnet sind.

Zur Einschätzung der curricularen Nähe und didaktische Relevanz wurde ein Lehrerfragebogen eingesetzt. Die Items wurden zum großen Teil als lehrplanrelevant, z. T. auch als didaktisch innovativ gemäß den aktuellen curricularen Entwicklungstendenzen bewertet.

In Einzel- und Gruppeninterviews mit Schülerinnen und Schülern (N=27) wurde anhand der Methode des *lauten Denkens* überprüft, mit welchen Strategien die Schülerinnen und Schüler die Aufgaben lösen. Problematische Items konnten so identifiziert werden und entfernt, bzw. korrigiert werden.

2.2.2 Theoretische Modelle

Zur Postulierung eines Kompetenzstrukturmodells werden verschiedene theoretische Modelle empirisch geprüft und miteinander verglichen. Die Repräsentationsformen Tabelle, Graf, und Situation (Wortvorschrift) bilden die Grundlage für mögliche Komponenten des Modells. Die Kompetenz im Umgang mit Termen wird nicht einbezogen, weil der systematische Erwerb dieser Fähigkeit erst mit der Klassenstufe 8 beginnt. Daher wird nicht erwartet, dass sich diese Kompetenzen in den untersuchten Klassenstufen 7 und 8 strukturiert in empirischen Daten widerspiegeln.

Modell 1: Repräsentationswechsel (4Dim)

Das Modell postuliert, dass die Kompetenz des Problemlösens mit funktionalen Repräsentationen durch die Fähigkeit der Übersetzung zwischen grafischer bzw. numerischer Darstellung und Situation sowie die Verarbeitung innerhalb der numerischen oder grafischen Repräsentationen bestimmt wird.

Beim Wechsel zwischen grafischer bzw. numerischer Repräsentation und Situation werden beide Übersetzungsrichtungen gleichzeitig modelliert. Bei den meisten Items war es nicht möglich, eindeutig eine Bearbeitungsrichtung anzunehmen. Eine solche Auflösung der kognitiven Prozesse durch Items erwies sich auf dem Komplexitätsniveau des Kompetenzkonstruktes als nicht möglich. Hieraus resultiert ein vierdimensionales Modell:

1. Komponente: Mathematisieren innerhalb der *grafischen* Darstellung (GG):

Verarbeiten grafischer Daten ohne Situationsbezug,

2. Komponente: Wechsel *situative* vs. *grafische* Repräsentation (SG):

Interpretation grafischer Daten, meist vorliegend als Funktionsgrafen,

3. Komponente: Mathematisieren innerhalb der *numerischen* Darstellung (NN):

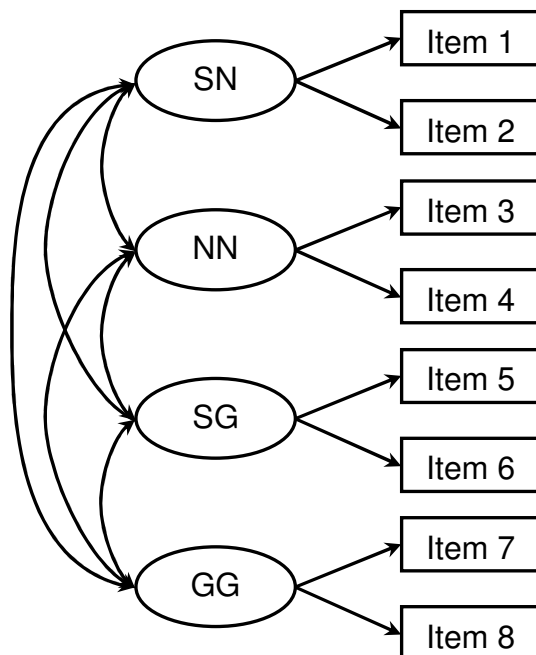
Verarbeitung numerischer Daten ohne Situationsbezug,

4. Komponente: Wechsel *situative* vs. *numerische* Repräsentation (SN):

Interpretation numerischer Daten, meist vorliegend in Tabellenform.

Eine schematische Darstellung dieses theoretischen Modells findet sich in Abbildung 7.

Abb. 7: Modell 1, Repräsentationswechsel

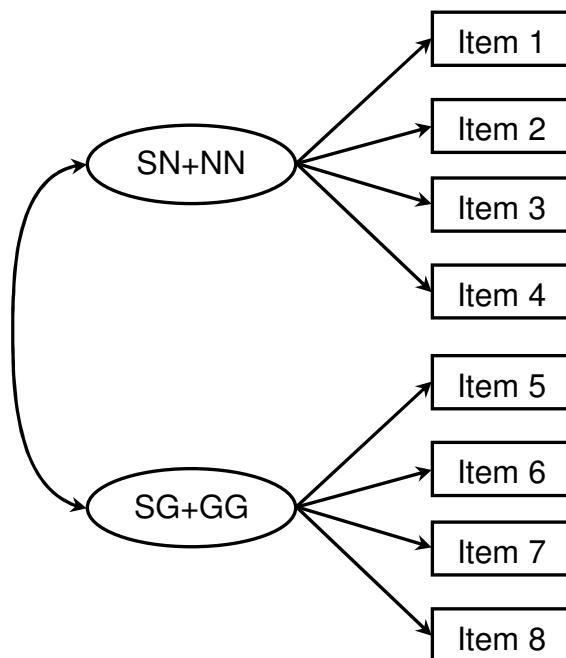


Anmerkung: SN= Situativ-numerisch, NN= Numerisch-numerisch, GG= Grafisch-grafisch, SG=Grafisch-situativ

Modell 2: Repräsentationstypen (2Dim)

Dieses Modell geht davon aus, dass die wesentliche Kompetenz zum Umgang mit funktionalen Repräsentationen durch zwei Komponenten bestimmt wird. Eine Komponente beinhaltet Aufgaben mit grafischem Inhalt, die andere umfasst numerische Aufgaben. Dabei spielt die Tatsache, ob eine Übersetzung zwischen Situation und grafischen, bzw. numerischen Daten stattfinden muss, keine entscheidende Rolle. Hieraus resultiert ein zweidimensionales Modell:

1. Komponente: Grafische Darstellung (GG und SG)
2. Komponente: Numerische Darstellung (NN und SN).

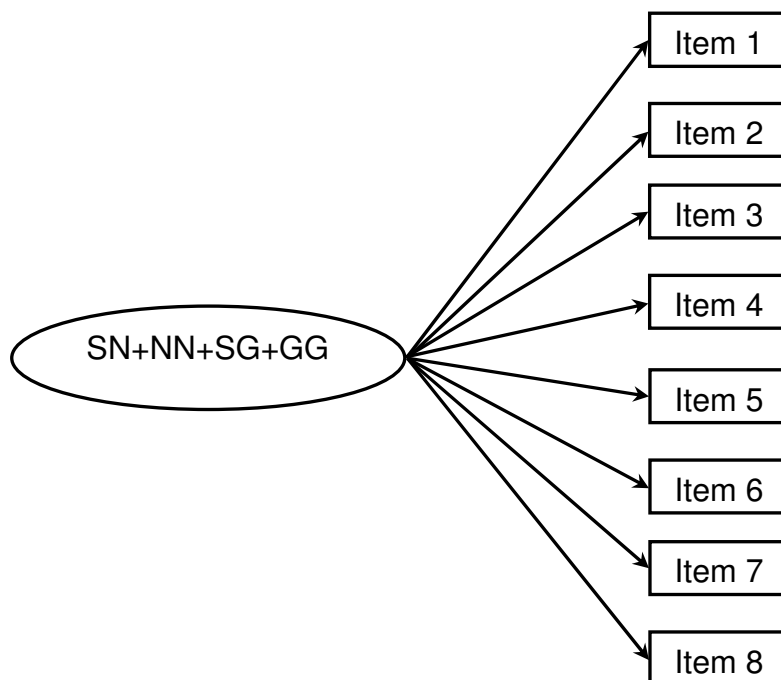
Abb. 8: Modell 2, Repräsentationstypen

Anmerkung: SN= Situativ-numerisch, NN= Numerisch-numerisch, GG= Grafisch-grafisch, SG=Grafisch-situativ

Modell 3: Repräsentationsunabhängige Kompetenz (1Dim)

Schließlich dient die Annahme, dass die Problemlöseleistung mit funktionalen Repräsentationen im wesentlichen als ein einziges Fähigkeitskonstrukt aufzufassen ist, als Vergleichsmodell:

Eine Komponente: GG, SG, SN und NN.

Abb. 9: Modell 3, Repräsentationsunabhängige Kompetenz

Anmerkung: SN= Situativ-numerisch, NN= Numerisch-numerisch, GG= Grafisch-grafisch, SG=Grafisch-situativ

2.2.3 Erhebungsdesign

Die Testung erfolgte anhand eines Multi-Matrix-Designs. Es wurden sieben Blöcke mit jeweils 11 Items gebildet, je drei Blöcke bilden ein Testheft. Die genaue Verteilung der Blöcke auf die einzelnen Testhefte ist Tabelle 1 zu entnehmen:

Tabelle 1: Multi-Matrix-Sampling-Design

Testheft	Testheft	Testheft	Testheft	Testheft	Testheft	Testheft
1	2	3	4	5	6	7
2	3	4	5	6	7	1
1	2	3	4	5	6	7
3	4	5	6	7	1	2

Von den insgesamt 80 Items des Mathematiktests wurde jedem Schüler ein Testheft mit 34 Items vorgelegt. Jedes Item wurde von ca. 370 Schülern bearbeitet. Als besondere Herausforderung stellte sich das Vorkommen von Parallelitems heraus, welche dieselbe Problemsituation enthalten, aber eine Lösung mit verschiedenen Repräsentationen vorgeben. Da die Lösung eines Problems in einer Repräsentation einen möglichen zweiten Lösungsversuch stark beeinflussen würde, durfte keines der Testhefte zwei Parallelitems enthalten. Dies konnte durch eine geeignete Verteilung paralleler Items auf Blöcke mit Mindestabstand von drei gewährleistet werden.

Die Testung fand während des Schulbetriebs und überwiegend im Mathematikunterricht statt. Die Tests wurden an zwei Terminen von geschulten Testleitern durchgeführt, die den Schülern in einer fünfminütigen Instruktion Art und Zweck der Untersuchung mitteilten. Für den Mathematiktest hatten die Schüler 75 Minuten Zeit.

2.2.4 Beschreibung der Stichprobe

Die Stichprobe umfasste 37 Gymnasialklassen, von denen 20 (54,1%) Klassen aus Baden-Württemberg und 17 (45,9%) Klassen aus Hessen stammen. 17 (45,9%) Klassen gehörten der siebten Klassenstufe an und 20 (54,1%) Klassen der achten Klassenstufe. Insgesamt nahmen 872 Schülerinnen und Schüler an der Studie teil, die Stichprobe setzte sich aus 471 (54,0%) Mädchen und 399 (46,0%) Jungen zusammen (2 Angaben fehlend).

Tabelle 2: Stichprobe

	Mädchen	Jungen	Gesamt
Klasse 7	185	216	401
Klasse 8	286	183	469
Gesamt	471	399	870

Es wurden ausschließlich Klassen aus Gymnasien berücksichtigt, um die Schülerpopulation möglichst homogen hinsichtlich sprachlicher Fähigkeiten und schulformtypischer Bildungsanforderungen zu halten. Der Einfluss potentieller Moderatorvariablen auf die Ausprägung differenzierterer kognitiver Strukturen soll so verringert werden.

2.2.5 Statistische Verfahren: Ein- und mehrdimensionale Raschmodelle

Mathematische Kompetenzen von Schülerinnen und Schülern, die durch psychometrische Skalen erfasst werden, haben den Status latenter Merkmale, deren Ausprägung durch die Messung manifester Indikatoren (z.B. Lösen von grafisch repräsentierten Aufgaben) geschätzt wird. Diese latenten Merkmale werden in der Testtheorie als erklärende Variablen aufgefasst, die das Antwortverhalten maßgeblich bedingen (Bollen 2002). Die Modellierung dieser Kompetenzen erfolgt mittels probabilistischer Testmodelle. Diese Modellierung hat sich in der Kompetenzdiagnostik insbesondere in den internationalen Vergleichsstudien TIMSS und PISA (Baumert et al., 2001; OECD 2005) als Standard etabliert, da sich eine Vielzahl diagnostischer Vorteile ergeben (z.B. Fairness der Vergleiche, spezifische Objektivität; metrische Kompetenzindizes; Bond & Fox, 2001; Rost, 2003). Die Basis bildet das Rasch Modell, das folgendermaßen beschrieben wird:

$$p(X_{vi} = 1) = \frac{\exp(\theta_v - \sigma_i)}{1 + \exp(\theta_v - \sigma_i)}$$

Um zu überprüfen, ob der mathematischen Fähigkeit, Probleme mit Funktionen zu lösen, mehrere latente Dimensionen zugrunde liegen, wurde das Multidimensional Random Coefficients Multinomial Logit Model (Adams, Wilson & Wang, 1997) zugrunde gelegt. Dieses ist bereits in mehreren großen Schulleistungsstudien zur Modellierung von Schülerkompetenzen zur Anwendung gekommen und ist in der Software ConQuest implementiert (Wu, Adams & Wilson, 2001). Das MRCMLM stellt eine mehrdimensionale Erweiterung des Rasch-Modells dar, anhand dessen ordinale und nominale beobachtete Variablen mit mehreren latenten Dimensionen analysiert werden können. Die Formel dieses Modells ist im Folgenden aufgeführt:

$$p(x_{ik} = 1; \mathbf{A}, \mathbf{B}, \boldsymbol{\delta} | \boldsymbol{\theta}) = \frac{\exp(\mathbf{b}_{ik} \boldsymbol{\theta} - \mathbf{a}'_{ik} \boldsymbol{\delta})}{\sum_{k=1}^{K_i} \exp(\mathbf{b}_{ik} \boldsymbol{\theta} - \mathbf{a}'_{ik} \boldsymbol{\delta})}$$

Beim Vergleich verschiedener Rasch-Modelle muss zur Beurteilung der Modellgüte abgewogen werden, wie gut die empirischen Daten zu dem Modell passen und wie viele Parameter für das Erreichen dieser Passung erforderlich waren (vgl. Rost 2004). Dazu werden u.a. informationstheoretische Indices herangezogen, die sowohl die Zahl der Modellparameter als auch die Likelihood als Anpassungsmaß berücksichtigen. Es handelt sich um das Consistent Akaike Information Criterion (AIC_c) und das Bayesian Information Criterion (BIC). Zusätzlich wird ein Signifikanztest durchgeführt. Es gilt sowohl für den AIC_c -Index als auch für den BIC, dass ein Modell um so besser passt, je geringer der Index ist. Nach Burnham und Anderson (2004) wird ein Modell dann als bedeutsam besser oder schlechter passend als ein anderes interpretiert, wenn die Differenz der Indices größer oder gleich zehn ist. Von Burnham und Anderson (2004) wird der AIC_c besonders dann empfohlen, wenn das Verhältnis von Stichprobengröße zu größter Parameterzahl der verglichenen Modelle geringer als 40 ist. Die Formeln dieser Indices befinden sich im nächsten Abschnitt:

$$AIC_c = -2 \log L + 2n_p + \frac{2n_p(n_p + 1)}{N - n_{p-1} - 1}$$

$$BIC = -2 \log L + (\log N) n_p$$

Zur Überprüfung der Itemqualität wurde die *weighted mean square statistics* (WMNSQ) herangezogen. Der WMNSQ repräsentiert die relative Diskrepanz in der Variation zwischen den empirischen Daten und der Modellvorhersage. Adams and Khoo (1996) schlagen einen Range von 0,75 bis 1,33 für den WMNSQ-Wert vor.

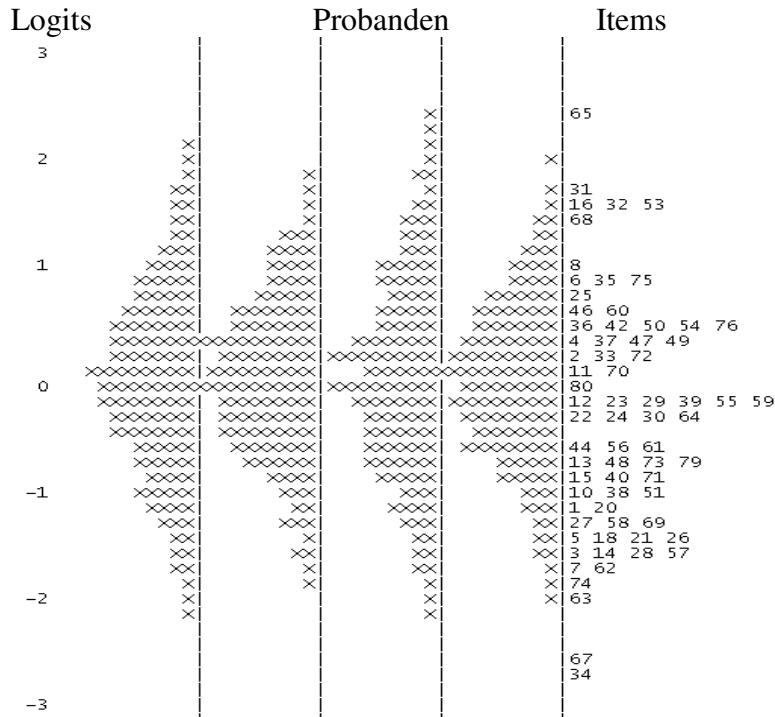
2.3 Ergebnisse

2.3.1 Deskriptive Ergebnisse

Die deskriptiven Statistiken für die verschiedenen theoretischen Modelle sind im Anhang unter Abschnitt 9.6 zusammengefasst. Die Itemschwierigkeit (δ) ist in Logits angegeben, welcher für den natürlichen Logarithmus des Lösen/Nicht-Lösen Verhältnisses steht. Zur Einschätzung des Itemfit ist der WMNSQ-Wert angegeben, alle Items weisen eine sehr gute Übereinstimmung mit dem Kriterium von Adams und Khoo (1996) auf.

Die geschätzten Itemschwierigkeiten in den verschiedenen Modellen sind vergleichbar, wie die deskriptive Analyse zeigt (vgl. 9.2). Die Item-Person-Map in Abb. 10 illustriert die Schätzungen der Personenparameter und Itemschwierigkeiten des vierdimensionalen Raschmodells. Diese bestätigt die gute Passung der Schwierigkeit der Items zum Leistungsniveau der Probandengruppe, da das Spektrum der Schülerkompetenzausprägungen durch die Itemschwierigkeiten vollständig abgedeckt wird.

Abb. 10: Item-Person-Map des Modells 1, Repräsentationswechsel



Anmerkung: Jedes „X“ repräsentiert 6,7 Probanden

Die Schwierigkeit der Items sollte zwischen -3 und 3 liegen (Baker, 2001), was von einem Item (# 9) nicht erfüllt wird. Alle anderen Items liegen in dem angegebenen Intervall.

2.3.2 Modellvergleich

Die empirische Überprüfung der Passung der verschiedenen Modelle, die in Absatz 2.2.2 aufgeführt sind, ergab das folgende Bild:

Tabelle 3: Modelfit

Modell	AIC _c	Differenz AIC _c	BIC	Differenz BIC	Parameter	-2 log L
Repräsentations- unabhängig (1Dim)	29985	-	30017	-	72	29908.9
Repräsentations- typen (2Dim: G vs N)	29975	-10	30007	-10	74	29898.4
Repräsentations- wechsel (4Dim:GG,SG,NN,S N)	29943	-32	29975	-32	81	29866.8

Anmerkung: Die Modelle sind vom höchsten bis niedrigsten AIC_c sortiert. Die Differenzen beziehen sich jeweils auf zwei aufeinanderfolgende Modelle. AIC_c und BIC sind gerundet. L = Likelihood

Das vierdimensionale Modell *Repräsentationswechsel* weist verglichen mit den anderen Modellen den geringsten AIC_c und BIC auf. Dieses Modell passt also auf der Grundlage dieses relativ strengen Kriteriums, das die Zahl der geschätzten Parameter, die Stichprobengröße und die Likelihood berücksichtigt, am besten zu den empirischen Daten.

Auch ein Chi²-Test ergab, dass das vierdimensionale Modell ($G^2(81) = 29866.83$) die Daten signifikant besser ($\Delta G^2(9) = 42,09$; $p < .001$) als das eindimensionale Modell ($G^2(72) = 29908.9$) erklärt. Das vierdimensionale Modell ($G^2(81) = 29866.8$) erklärt die Daten ebenfalls signifikant besser ($\Delta G^2(7) = 31,56$; $p < .001$) als das zweidimensionale Modell ($G^2(74) = 29898,3$). Die AIC_c-Differenz von vierdimensionalen und eindimensionalen Modell beträgt 42, das 1DIM-Modell und das 2Dim-Modell weist eine AIC_c-Differenz von 32 auf. Das

Modell 4Dim repräsentiert damit die empirischen Daten am besten, was die zugrundeliegende Theorie der Repräsentationswechsel stützt.

2.3.3 Überprüfung der Skalenqualität: Reliabilitäts- und Korrelationsanalyse

Zur Absicherung der Modellstruktur wurden die Reliabilitäten der einzelnen Komponenten der verschiedenen Modelle und die Korrelationen zwischen den Dimensionen bestimmt. Zu diesem Zweck wurde die EAP/PV-Reliabilität eingesetzt, die im Rahmen der Rasch-Analyse bestimmt wurde und mit Cronbachs Alpha vergleichbar ist (Rost, 2004).

In Tabelle 4 sind die latenten Korrelationen zwischen den Komponenten des vierdimensionalen Modells aufgeführt, die EAP/PV-Reliabilitäten befinden sich in der Diagonale, in Klammern steht die Anzahl der Items in der jeweiligen Dimension.

Tabelle 4:
Latente Korrelationen und Reliabilitäten des Modells *Repräsentationswechsel* (4Dim)

	SN	NN	GG	SG
SN (15)	0.64			
NN (05)	0.65	0.62		
GG (18)	0.70	0.57	0.52	
SG (32)	0.76	0.71	0.64	0.72

Anmerkung: Die Reliabilitätskoeffizienten sind in der Diagonale aufgeführt.

Die Reliabilität der Dimension GS ist mit einer EAP/PV-Reliabilität von .72 als zufriedenstellend anzusehen. Die niedrigen Reliabilitäten der anderen Skalen sind teilweise durch geringe Itemzahlen zu erklären. Hervorzuheben sind die vergleichsweise moderaten latenten Korrelationen zwischen den Skalen des vierdimensionalen Modells. Dieser Befund legt nahe, dass es gelingt, durch den entwickelten Itempool trennbare Kompetenzaspekte zu erfassen.

Das eindimensionale Modell weist eine zufrieden stellende EAP/PV-Reliabilität von $r=0,76$ auf. Die EAP/PV-Reliabilität der Komponente Grafische Darstellung (GG und SG) des zweidimensionalen Modells beträgt $r=0,69$ und die EAP/PV-Reliabilität der Komponente Numerische Darstellung (NN und SN) liegt bei $r=0,65$. Die latente Korrelation der beiden Komponenten liegt bei $r=.78$.

2.4 Diskussion und Ausblick

Unsere Befunde geben Hinweise darauf, dass der Repräsentationstyp, insbesondere der Wechsel der Repräsentation für den mathematischen Inhaltsbereich „funktionale Veränderung“ bedeutsam für die Kompetenzstruktur ist. Es wurde ein vierdimensionales Kompetenzstrukturmodell vorgeschlagen, welches die grafische und numerische Modellierung von Situationen bzw. den Wechsel zwischen diesen Repräsentationen beinhaltet. Dieses Modell ermöglicht die empirisch beste Beschreibung des Kompetenzbereichs.

Die Analyse der Binnenstruktur dieses Modells zeigt allerdings noch unbefriedigende Reliabilitäten für drei der vier Dimensionen. Die latenten Korrelationen der vier Dimensionen des Kompetenzstrukturmodells sind erwartungsgemäß hoch, aber noch akzeptabel. Es wurden in der PISA-Studie höhere latente Korrelationen zwischen verschiedenen Domänen gemessen (OECD, 2005). Der Itempool soll bezüglich der identifizierten Dimensionen optimiert werden, um eine höhere Konstruktvalidität und die Reliabilität der Skalen und damit eine bessere Einsetzbarkeit als Diagnostikum zu gewährleisten.

Eine itembasierte Kompetenzmessung kann die kognitiven Prozesse, die dem modellierten Fähigkeitsbereich zugrunde liegen, immer nur indirekt messen. Selbst bei Items, die auf eine Repräsentation fokussieren, kann nicht immer ausgeschlossen werden, dass Schülerinnen und Schüler andere Repräsentationen hinzuziehen. Dieses könnte durch eine experimentelle Studie kontrolliert werden, indem die Probanden etwa beim Lösen einer Aufgabe unter Zeitdruck gesetzt werden. Dadurch hätten die Schüler keine Zeit, zusätzlich ihren eigenen Lösungsweg zu der in der Aufgabenstellung geforderten Repräsentationsform zu verwenden.

Die identifizierten vier Dimensionen des Modells sollen empirisch in Kompetenzniveaus unterteilt werden. Dadurch wird es möglich, zwischen Personen mit unterschiedlichen Teilkompetenzen und Teilkompetenzprofilen zu unterscheiden und es können Kompetenzentwicklungen abgebildet werden.

Das postulierte Kompetenzstrukturmodell kann die Basis für ein Diagnoseinstrument mathematischer Problemlösefähigkeit beim Umgang mit Funktionen bilden und gleichzeitig Ansatzpunkte für die Förderung bieten. Diese kann sowohl zur Analyse von Entwicklungsverläufen in Längsschnittstudien dienen als auch in der Schulpraxis eingesetzt werden. Ziel des Diagnostikums ist es, Erkenntnisse zu Förderbedarf und Förderungsmöglichkeiten auf Lerngruppen- und Individuenebene zu gewinnen. Die externen Repräsentationen von Funktionen z.B. als Tabelle oder als Graf sind in Deutschland ab Klassenstufe 5 relevant, in der 8. Klasse werden auch Terme eingeführt. Daher kann ein Kompetenzmodell auf Basis der verschiedenen Repräsentationsformen von Funktionen von der 5. bis zur 10. Klassenstufe eingesetzt werden.

3 STUDIE 2:

Modellierung mathematischer Kompetenz anhand von

Within-Item-Multidimensionality

3.1 Einleitung

Zur Modellierung mathematischer Kompetenz beim Problemlösen mit Funktionen wurde auf der Basis von verschiedenen Repräsentationsformen und Repräsentationswechsel (z.B. Swan, 1985; Ainsworth et al., 2002; Goldin, 1998) ein vierdimensionales Kompetenzstrukturmodell postuliert (siehe Kapitel 2). In diesem Modell wird angenommen, dass die Kompetenz des Problemlösens mit Funktionen durch die spezifischen Wechsel zwischen Repräsentationstypen bestimmt wird. Die Fähigkeit der Übersetzung zwischen grafischer bzw. numerischer Darstellung und Situation sowie die Verarbeitung innerhalb der Repräsentationen werden als psychometrisch trennbare Komponenten aufgefasst. Im Folgenden soll überprüft werden, ob sich die Kompetenzstruktur im mathematischen Inhaltsbereich „Wachstum und Veränderung“ anhand eines alternativen multidimensionalen Modells besser erklären lässt.

Wenn multidimensionale Modelle angewendet werden, ist zu unterscheiden, ob jedes Item als Indikator jeweils nur für eine latente Variable eingesetzt wird, oder jede Aufgabe durch mehrere latente Variablen bestimmt wird. Der erste Fall entspricht einer Einfachstruktur der Faktorladungen wie in der Faktorenanalyse, der zweite beinhaltet ein Ladungsmuster mit Mehrfachladungen. Man spricht im Kontext von IRT-Modellen von

„Between-Item-Multidimensionality“ und „Within-Item-Multidimensionality“ (Adams et al., 1997).

Eine itembasierte Kompetenzmessung kann die kognitiven Prozesse, die dem modellierten Fähigkeitsbereich zugrunde liegen, immer nur indirekt messen. Selbst bei Items, die auf eine Repräsentation fokussieren, kann nicht immer ausgeschlossen werden, dass Schülerinnen und Schüler andere Repräsentationen hinzuziehen. Daher soll in der vorliegenden Studie versucht werden, anhand der Lösungswege, welche die Schüler/innen auf dem Testbogen angegeben haben, diese kognitiven Prozesse genauer abzubilden. Der Test beinhaltet Aufgaben, in denen die Schüler/innen aufgefordert werden, eine bestimmte Repräsentationsart zu Lösung des Items zu verwenden.

Aufgaben, die von den Probanden z.B. sowohl anhand eines Grafen als auch mit einer Tabelle gelöst werden können, sollen –falls sich dieses in den empirischen Daten zeigt– mehreren Dimensionen des Kompetenzmodells zugewiesen werden („Within-Item-Multidimensionality“, Adams et al., 1997). Hierbei wird nicht erwartet, dass die Schülerinnen und Schüler ausschließlich diese Repräsentationsform zur Lösung des Items verwenden. Wenn in der Aufgabe eine tabellarische Lösung gefordert ist, die Schüler jedoch einen Grafen verwenden, kann eine Übersetzung zwischen Graf und Tabelle stattgefunden haben, so dass sie möglicherweise beide Repräsentationsformen genutzt haben. Insofern kann anhand der Notizen der Probanden nicht unmittelbar auf kognitive Prozesse geschlossen werden. Diese werden lediglich als Hilfsmittel verwendet, um kognitive Prozesse abzubilden.

Für die vorliegende Studie ergibt sich die folgende Fragestellung:

Kann ein Kompetenzmodell welches die Lösungswege der Schüler/innen berücksichtigt, die empirischen Daten besser erklären als das in Kapitel 2 postulierte Kompetenzstrukturmodell („*Between-Item-Multidimensionality*“)?

3.2 Hypothese

Es wird erwartet, dass ein *Within-Item-Multidimensionality*-Modell, das die Lösungswege der Schüler/innen berücksichtigt, die empirischen Daten besser erklären kann, als ein vierdimensionales *Between-Item-Multidimensionality*-Modell (vgl. Kapitel 2).

3.3 Methode

Die Methode, d.h. die Itemkonstruktion, die Beschreibung der Stichprobe und das Testheftdesign sind in Kapitel 2.2 beschrieben.

3.3.1 Theoretische Modelle

Zusätzlich zu den in Abschnitt 2.4 postulierten Modellen, einem ein-, einem zwei- und einem vierdimensionalen *between-item-multidimensionality*-Modell, werden in der vorliegenden Studie zwei weitere Modelle hinsichtlich ihrer empirischen Passung überprüft. Die Modelle 1-3 sind in Kapitel 2.2.2 beschrieben.

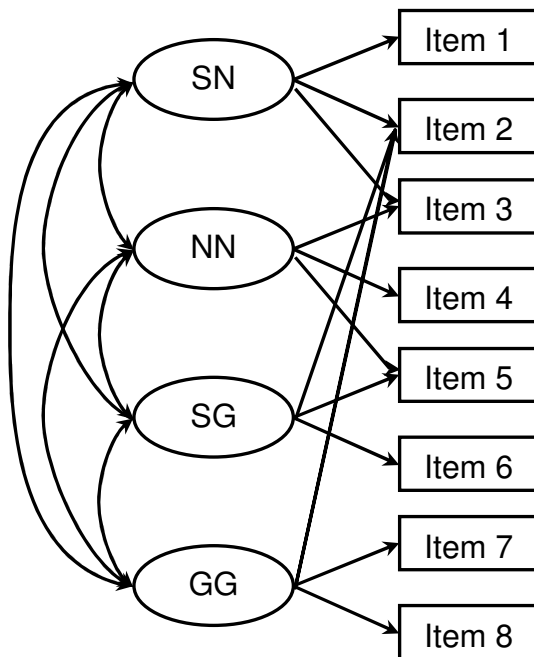
Modell 4: Repräsentationswechsel (4Dim-*within*)

In diesem Modell wird angenommen, dass die Kompetenz des Problemlösens mit funktionalen Repräsentationen durch die spezifischen Wechsel zwischen Repräsentationstypen bestimmt wird. Es handelt sich um ein *Within-Item-Multidimensionality*-Modell, so dass ein Item zu mehreren Dimensionen gehören kann, wenn dies aufgrund der Auswertung der Lösungswege plausibel ist.

1. Komponente: Mathematisieren innerhalb der *grafischen* Darstellung (GG):
Verarbeiten grafischer Daten ohne Situationsbezug,
2. Komponente: Wechsel *situative* vs. *grafische* Repräsentation (SG):
Interpretation grafischer Daten, meist vorliegend als Funktionsgraf,
3. Komponente: Mathematisieren innerhalb der *numerischen* Darstellung (NN):
Verarbeitung numerischer Daten ohne Situationsbezug
4. Komponente: Wechsel *situative* vs. *numerische* Repräsentation (SN):
Interpretation numerischer Daten, meist vorliegend in Tabellenform.

Eine schematische Darstellung dieses theoretischen Modells findet sich in der folgenden Abbildung 11:

Abb.11: Modell 4 Repräsentationswechsel (4Dim-*within*)



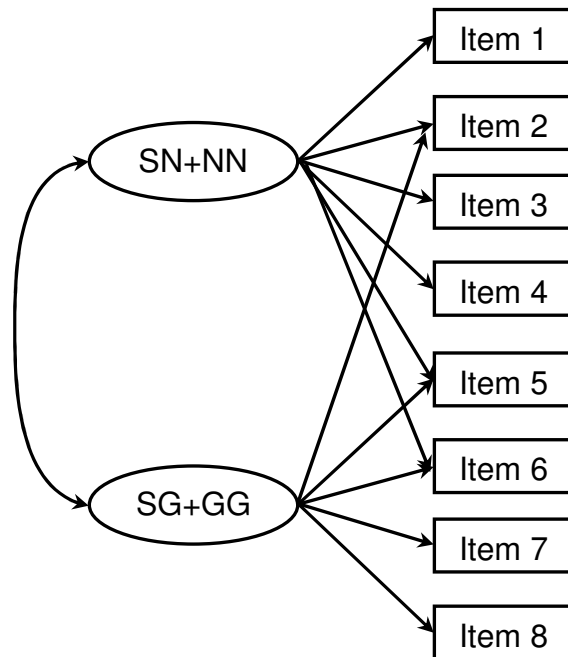
Modell 5: Repräsentationstypen (2Dim-*within*)

Dieses Modell geht davon aus, dass die wesentlichen, unterscheidbaren Komponenten der Kompetenz des Problemlösens mit funktionalen Repräsentationen durch den Repräsentationstyp bestimmt sind. Auch in diesem Modell ist es möglich, dass Items zu beiden Dimensionen gehören.

1. Komponente: Grafische Darstellung (GG und SG)
2. Komponente: Numerische Darstellung (NN und SN).

Eine schematische Darstellung dieses Modells ist in Abb. 12 veranschaulicht.

Abb.12: Modell 5 Repräsentationstypen (2Dim-*within*)



3.3.2 Statistisches Verfahren: Within-Item-Multidimensionality

Zur Modellierung von komplexen Zusammenhängen, bei denen Aufgaben gleichzeitig durch mehrere latente Variablen beeinflusst werden, verwendet man in der IRT „*Within-Item Multidimensionality*“-Modelle (Adams et al., 1997). Diese sind in z.B. der Software ConQuest (Wu et al., 1998) oder Mplus (Muthen & Muthen, 2006) implementiert.

Man benötigt theoretisch fundierte a-priori-Annahmen über Merkmale der Items und die inhaltliche Bedeutung der unterschiedlichen latenten Variablen, um eine solche Struktur zu definieren. In der vorliegenden Studie werden die in den Testheften angegebenen Lösungswege der Probanden benutzt, um auf kognitive Prozesse zu schließen. Walker und

Beretvas (2003) beschreiben ein mehrdimensionales IRT-Modell, welches als Beispiel für ein theoretisch fundiertes Modell mit Mehrfachladungen dienen kann. Die Autorinnen postulieren ein zweidimensionales Modell zur Erfassung mathematischer Kompetenz, wobei alle Items als Indikatoren für eine allgemeine mathematische Fähigkeit dienen. Zusätzlich wurde ein zweiter Faktor angenommen, auf dem Items mit offenem Antwortformat außerdem laden, welcher als die Fähigkeit zur Kommunikation mathematischer Inhalte interpretiert wurde.

Die einzelnen Items können also auf jeweils unterschiedlich vielen Dimensionen laden. Um beispielsweise ein Item mit einem funktionalen Zusammenhang zu lösen, können Schüler/innen sowohl einen Grafen als auch eine Tabelle verwenden. Dieses Item würde in einem Modell mit „*Within-Item Multidimensionality*“ zwei latenten Dimensionen zugeordnet.

3.4 Ergebnisse

3.4.1 Deskriptive Analyse der Lösungswege

In der Auswertung der Items des Mathematiktests wurden bei insgesamt 18 Items die Lösungswege mit berücksichtigt. In der folgenden Tabelle finden sich die deskriptiven Ergebnisse der Auswertung der Lösungswege. Es ist für jedes der Items angegeben, wie viel Prozent der Probanden es anhand eines Schaubildes oder mit Hilfe einer Tabelle, bzw. Rechnung gelöst haben, in Klammern finden sich die absoluten Zahlen. Zusätzlich sind der prozentuale Anteil und die absolute Zahl von Schüler/innen angegeben, die das jeweilige Item nicht bearbeitet haben, obwohl es ihnen vorgelegt wurde.

Tabelle 5: Deskriptive Analyse der Lösungswege

Aufgabentyp	Item	Schaubild	Tabelle/ Rechnung	Fehlend
Offene Aufgabe	78_150aSxGN	25 (88)	36,36 (128)	38,64(136)
	66_158Sxgn	32,87 (118)	43,73 (157)	23,40 (84)
	50_26bSNB	13,59 (50)	43, 48 (160)	42,93 (158)
	43_150bSxGN	13,26 (50)	37,93 (143)	48,81 (184)
	41_30cSGB	7,07 (27)	74,87 (286)	18,06 (69)
	19_151SxGN	15,79 (57)	72,02 (260)	12,19 (44)
	17_26aSNB	9,37 (34)	69,15 (251)	21,49 (78)
Lösung soll anhand eines Schaubildes erfolgen	74_30aSGB	92,90 (327)	1,99 (7)	5,11 (18)
	32_72aSGB	69,92 (272)	8,23 (32)	21,85 (85)
	9_35dSGB	63,10 (224)	6,48 (23)	30,42 (108)
Lösung soll anhand einer Tabelle oder Rechnung erfolgen	65_72bSNB	7,2 (26)	75,9 (274)	16,9 (61)
	5_30bSNB	7,4 (27)	72,05 (263)	20,55 (75)
Graf Ablesen	75_16aGGB	29,75 (105)	37,11 (131)	33,14 (117)
	20_14aGGB	8,63 (32)	28,03 (104)	63,34 (235)
Zuordnung proportional	62_31aNNC	23,06 (83)	36,67 (132)	40, 28 (145)
Tabelle ablesen	49_16bNNB	27,99 (103)	54,89 (202)	17,12 (63)
	47_5cSNB	11,92 (44)	22,49 (83)	65,58 (242)
	22_7NNC	9,51 (37)	50,09 (198)	39,59 (154)

Anmerkung: Die Zahlen sind in Prozent aufgeführt, in Klammern sind die absoluten Zahlen der Schülerinnen und Schüler angegeben, die einen bestimmten Lösungsweg gewählt haben.

3.4.2 Zuordnung der Items zu Dimensionen

Zur Beantwortung der Forschungsfrage wurden die Items auf der Basis der in 4.4.1 dargestellten deskriptiven Analysen der Lösungswege den verschiedenen Dimensionen des Within-Item-Multidimensionality-Modells zugeordnet. Kriterium für die Zuordnung eines Items zu einer Dimension war, dass mindestens 10 % der Probanden das Item entsprechend gelöst haben. Insgesamt wurden 6 Items (Item 50_26bSNB, Item 75_16aGGB, Item

20_14aGGB, Item 62_31aNNC, Item 49_16bNNB, Item 47_5cSNB) zusätzlich einer weiteren Dimension zugeordnet (vgl. Tabelle 6).

Tabelle 6: Zuordnung der Items zu Dimensionen

Aufgabentyp	Item	4Dim (between)	4Dim (within)	2Dim (between)	2Dim (within)
Offene Aufgabe	78_150aSxGN	-	-		
	66_158Sxgn	-	-		
	50_26bSNB	SN	SN+SG	N	NG
	43_150bSxGN	-	-		
	41_30cSGB	-	-		
	19_151SxGN	-	-		
	17_26aSNB	-	-		
Lösung soll anhand eines Schaubildes erfolgen	74_30aSGB	SG	SG	G	G
	32_72aSGB	SG	SG	G	G
	9_35dSGB	SG	SG	G	G
Lösung soll anhand einer Tabelle oder Rechnung erfolgen	65_72bSNB	SN	SN	N	N
	5_30bSNB	SN	SN	N	N
Graf Ablesen	75_16aGGB	GG	GG+NN	G	NG
	20_14aGGB	GG	GG+NN	G	NG
Zuordnung proportional	62_31aNNC	NN	NN+GG	N	NG
Tabelle ablesen	49_16bNNB	NN	NN+GG	N	NG
	47_5cSNB	SN	SN+GG	N	NG
	22_7NNC	NN	NN	N	N

3.4.3 Modellvergleich

Die empirische Überprüfung der Passung der verschiedenen Modelle, die in Absatz 4.3.1 aufgeführt sind, ergab das folgende Bild:

Tabelle 7: Modellvergleich

Modell	AIC _c	Dif- ferenz AIC _c	BIC	Dif- ferenz BIC	Para- meter	-2 log L
Repräsentations- unabhängig (1D)	29985	-	30017	-	72	29908.91
Repräsentations- typen (2D: G vs N) <i>Between Item Dimensionality</i>	29975	-10	30007	-10	74	29898.38
Repräsentations- wechsel (4D:GG,SG,NN,SN) <i>Between Item Dimensionality</i>	29943	-42	29975	-42	81	29866.82
Repräsentations- typen (2D: G vs N) <i>Within Item Dimensionality</i>	30164	179	30220	203	74	30002.31
Repräsentations- wechsel (4D:GG,SG,NN,SN) <i>Within Item Dimensionality</i>	30236	251	30295	278	81	30057.35

Hinweis: Die Differenzen beziehen sich jeweils auf das eindimensionale Modell. AIC_c und BIC sind gerundet. L = Likelihood

Beim Vergleich der verschiedenen Modelle erzielte das 4Dim-*between*-Modell die beste empirische Passung (AIC_c = 29943, BIC=29975). Sowohl das 4Dim-*within*-Modell (AIC_c =30236) als auch das 2Dim-*within*-Modell (AIC_c = 30164) erklären die empirischen Daten deutlich schlechter. Es gilt sowohl für den AIC_c- Index als auch für den BIC, dass ein

Modell um so besser passt, je geringer der Index ist. Nach Burnham und Anderson (2004) wird ein Modell dann als bedeutsam besser oder schlechter passend als ein anderes interpretiert, wenn die Differenz der Indices größer oder gleich zehn ist. Das relativ strenge Kriterium des Consistent Akaike Information Criterion (AIC_c) berücksichtigt sowohl die Zahl der Modellparameter als auch die Likelihood als Anpassungsmaß. Auch der BIC (Bayesian Information Criterion) und ein Signifikanztest ($\Delta G^2(9) = 42,09; p < .001$) weisen das 4Dim-*between*-Modell als das empirisch am besten passende Modell aus.

3.5 Diskussion

Die Hypothese, dass ein Kompetenzmodell, in dem die Items auf mehreren Dimensionen laden („Within-Item-Multidimensionality“), die empirischen Daten besser erklären kann als in einem Modell mit Einfachstruktur, konnte nicht bestätigt werden. Das in Kapitel 2 beschriebene vierdimensionale Kompetenzstrukturmodell („Between-Item-Multidimensionality“) war den anderen geprüften Modellen bezüglich der Passung zu den empirischen Daten überlegen (vgl. Hartig & Höhler, 2008).

Möglicherweise wird durch die Interkorrelationen zwischen den Dimensionen verhindert, dass die Modellgüte verbessert wird. Daher verbessern die zusätzlichen Pfade, wie sie in dem Modell zur *Within-Item-Multidimensionality* angenommen werden, die Erklärung der empirischen Daten durch das Modell nicht. Würden die Dimensionen stärker trennen, wäre wahrscheinlich eine Verbesserung des *between*-Modells durch Mehrfachladungen der Items zu erwarten.

Eine andere Möglichkeit zur Erklärung, warum das Modell mit Mehrfachstruktur (*Within-Item-Multidimensionality*) die Daten nicht besser erklären kann ist, dass die

kognitiven Prozesse durch die Lösungswege der Schülerinnen und Schüler durch ein *Within-Item-Multidimensionality*-Modell nicht angemessen abgebildet werden. Diese könnten alternativ anhand von kognitiven Diagnosemodellen wie DINO oder DINA (z.B. Rupp, 2005) modelliert werden. Voraussetzung ist jedoch eine genaue vorherige Analyse der kognitiven Prozesse, die beim Lösen der Items eine Rolle spielen, beispielsweise durch die Methoden des lauten Denkens.

Insgesamt zeigt die Studie, wie schwer kognitive Prozesse bei Items zu modellieren sind, die anhand verschiedener Repräsentationsformen, wie z.B. der tabellarischen oder grafischen Repräsentation gelöst werden können. Daher wäre es empfehlenswert bei einer Optimierung des Itempools verstärkt fokussierte Items zu konstruieren, die nur eine Dimension abbilden.

4 STUDIE 3:

Repräsentationswechsel beim Umgang mit Funktionen – Identifikation von Kompetenzprofilen auf der Basis eines Kompetenzstrukturmodells

4.1 Theoretischer Hintergrund

Auf der Basis des in Kapitel 2 identifizierten Kompetenzstrukturmodells sollen typische Kompetenzprofile von Schülerinnen und Schülern beim Arbeiten mit mathematischen Funktionen und ihren Repräsentationsformen identifiziert werden. Trotz der hohen Relevanz innerhalb der Mathematikdidaktik, steht eine systematische empirische Untersuchung des Wechsels zwischen „Tabelle“ bzw. „Term“ und anderen externen Repräsentationsformen von Funktionen noch weitgehend aus.

Empirische Untersuchungen zum Multimedia-Lernen zeigen, dass unterschiedliche Repräsentationsformen mathematischer Zusammenhänge und der Wechsel zwischen diesen entscheidend für individuelle Lernprozesse sind. Damit sich ein positiver Lerneffekt einstellt, müssen die Schülerinnen und Schüler in der Lage sein, zwischen den vorliegenden Repräsentationen zu wechseln, die verbindenden Elemente zu identifizieren und zu verknüpfen (vgl. Ainsworth et al., 2002; Seufert 2003). Während des Erlernens dieser Grundfertigkeiten zeigen sich typische Fehler (vgl. Bodemer et al., 2004). Kozma und Russell (1997) berichten, dass Novizen beispielsweise verschiedene Repräsentationen nur über deren Oberflächenmerkmale aufeinander beziehen, Expertinnen und Experten hingegen stellen Bezüge über dahinterliegende Konzepte her.

Eine Möglichkeit der Beschreibung kognitiver Prozesse bei der Entnahme von Informationen aus Texten und Grafen wird im Modell zur integrativen Bild- und Textverarbeitung (vgl. Schnotz, 2005) beschrieben. Dieses Modell beschäftigt sich damit, wie Informationen, die aus Texten und Grafen entnommen wurden, zueinander in Beziehung gesetzt werden. Man unterscheidet *deskriptionale Repräsentationen*, zu denen Texte gehören, und *depiktionale Repräsentationen*, zu denen Grafen zählen. Die Prozesse des Bild- und Textverstehens greifen bei der integrativen Verarbeitung von Bild und Text ineinander und ergänzen sich auf der einen Seite inhaltlich, sie können den Interpretationsspielraum der jeweils anderen Repräsentationsform aber auch einschränken (Schnotz 2005). Es werden die folgenden Verarbeitungsstufen der Bild-Text-Integration unterschieden: Zunächst werden Detailinformationen aus Bild und Text miteinander verknüpft. Danach werden einfache semantische Relationen zwischen Text und Bild erkannt und auf die jeweils andere Repräsentationsform bezogen. Schließlich werden komplexe Relationsgefüge von Text und Bild verknüpft. Schnotz (2005) geht davon aus, dass sogenannte „graf schema“, also im Langzeitgedächtnis gespeicherte kognitive Schemata, existieren und für die Verarbeitung von Grafen von Bedeutung sind. Sie enthalten Informationen darüber, wie Informationen in Grafen vorliegen und wie sie entnommen werden können.

Solche Kompetenzen von Schülerinnen und Schülern hinsichtlich der Kombination von Graf bzw. Tabelle und Text sind z.B. für das Problemlösen mit Funktionen im Mathematikunterricht von Bedeutung. Unterschiedliche Schülerprofile beim Umgang mit Funktionen können auf verschiedene Verarbeitungstypen und -niveaus bei der Integration von Bild und Text zurückgeführt werden.

Pesonen, Ehmke und Haapasalo (2005) untersuchten den Umgang mit verschiedenen Repräsentationsformen im Themengebiet der binären Operationen an einer Stichprobe von

Studierenden. Es ließen sich anhand einer latenten Klassenanalyse drei qualitativ verschiedene Verständnisstufen vom mathematischen Begriff der binären Operation nachweisen.

Ein Cluster von Studierenden, das in allen untersuchten Inhaltsbereichen am besten abschnitt, konnte inhaltsgleiche Darstellungen zuordnen, binäre Operationen in verschiedenen Repräsentationsformen identifizieren, sowie symbolische, verbale oder grafische Beispiele produzieren. Ein anderes Cluster von Studierenden konnte zwischen verschiedenen Repräsentationsformen wechseln, wies aber Defizite in der Identifikation von verschiedenen Repräsentationen auf. Schließlich ergab sich ein drittes Cluster, das in allen drei untersuchten Bereichen die geringsten Werte erreichte. Aufgrund dieser Ergebnisse erscheint es plausibel, die Fähigkeit des Repräsentationswechsels psychometrisch differenzierter zu modellieren.

Ob und inwiefern bestimmte Kompetenzprofile mit kognitiven Leistungsaspekten zusammenhängen, wurde bisher noch nicht untersucht. Allerdings zeigen Befunde zur Präferenz von medial präsentierten Lernumgebungen, dass schwächere Schülerinnen und Schüler von grafischer Repräsentation einer Aufgabe profitieren (vgl. Snow & Yalow, 1982).

4.2 Ziele der Studie

Ausgehend von dieser Literaturlage erscheint es plausibel, von differenzierten Kompetenzprofilen beim Umgang mit mathematischen Repräsentationsformen auszugehen. Ziel der vorliegenden Studie ist es, auf Grundlage des beschriebenen Kompetenzmodells (vgl. Kapitel 2). Kompetenzprofile von Schülerinnen und Schülern der 7. und 8. Klasse abzuleiten und den Zusammenhang dieser Profile mit der Fähigkeit im figuralen Denken – als einem zentralen Aspekt der kognitiven Leistungsfähigkeit – zu untersuchen. Man versteht darunter die Fähigkeit, figural-räumlich präsentierte Probleme zu lösen (vgl. Heller & Perleth, 2000). Die Hypothese ist, dass eine hohe Kompetenz im figuralen Denken mit einer hohen Leistung bei Aufgaben einhergeht, die grafisch repräsentiert sind¹.

4.3 Methode

4.3.1 Erhebungsdesign

Das Erhebungsdesign sowie das Kompetenzmodell, das für die vorliegende Studie zugrunde gelegt wird, sind in Studie 1, Kapitel 2.2.1 - 2.2.3 dargestellt.

¹ Andere mögliche Moderatorvariablen wie Motivation, Mathematikangst oder selbstreflektiertes Lernen wurden nicht analysiert, da sich kein korrelativer Zusammenhang mit der Leistung im Mathematiktest ergeben hat.

4.3.2 Statistisches Analyseverfahren: Latente Klassenanalyse

Als geeignetes Verfahren zur Analyse von Kompetenzprofilen wurde eine Latente Klassenanalyse (LCA) durchgeführt. Der LCA liegt ein psychometrisches Modell aus der probabilistischen Testtheorie zugrunde: Es wird angenommen, dass die Zugehörigkeit einer Schülerin oder eines Schülers zu einer latenten qualitativen Merkmalsklasse ausreichend ist, um deren oder dessen manifeste Merkmalsausprägungen bis auf eine stochastische Komponente vorherzusagen. Das Modell postuliert, dass jede/r Schüler/in einer Klasse angehört, die sich durch ein spezifisches Fähigkeitsprofil auf den Modelldimensionen auszeichnet. Für jede Klasse werden klassenspezifische Erwartungswerte auf den vier Analyseskalen angenommen. Ein solches Modell hat im Rahmen einer Kompetenzdiagnostik den Vorteil, dass Personen nach Fähigkeitsprofilen unterschieden werden können. Die Feststellung individueller Stärken oder Schwächen ist dann geeigneter Ausgangspunkt einer Fördermaßnahme (vgl. Hartig, 2007).

Die LCA wurde mit der Software Latent Gold 4.5 durchgeführt. In einem iterativen Verfahren wurde dabei für post hoc vorgegebene Klassenzahlen nach dem Maximum-Likelihood-Prinzip die optimale Lösung erstellt. Die Entscheidung bei der Auswahl des Lösungsmodells wird anhand des CAIC getroffen (vgl. Vermunt 2004). Die durch die Erhebung in Schulklassen bedingte Mehrebenenstruktur wurde durch den in Latent Gold 4.5 (GClasses, vgl. Vermunt 2008) implementierten Expectation-Maximization-Algorithmus berücksichtigt.

4.3.3 Zusammenhang von Kompetenzprofilen und figuralem Denken

Um erste Hinweise auf Erklärungen für Fähigkeitsprofile zu erhalten, wurde das figurale Denken als Kovariate ausgewertet. Es wird vermutet, dass eine hohe Kompetenz im figuralen Denken mit einer hohen Leistung bei den grafisch repräsentierten Aufgaben einhergeht.

Das figurale Denken wurde mit dem Untertest „Figurenalogien“ (N2) aus dem Kognitiven Fähigkeitstest (Heller & Perleth, 2000) erhoben und ist in T-Werten ($M=50$, $s=10$) angegeben. Die Subskala N 2 besteht aus 25 figuralen Items, auf deren Basis ein Fähigkeitswert bestimmt wurde (Cronbachs Alpha = 0,80). Um zu überprüfen, ob sich die identifizierten Cluster von Schülerinnen und Schülern mit bestimmten Kompetenzprofilen in der Ausprägung des figuralen Denkens unterscheiden, wurde eine univariate Varianzanalyse (ANOVA) durchgeführt.

4.4 Ergebnisse

4.4.1 Analyse von Kompetenzprofilen

Zur Überprüfung, ob es Schülerinnen und Schüler gibt, die bezüglich der postulierten vier Dimensionen unterschiedliche Fähigkeitsmuster aufweisen, wurde eine latente Klassenanalyse (LCA) unter Berücksichtigung der Mehrebenenstruktur durchgeführt.

In der 7. Klassenstufe ließen sich keine typischen Profile von Schülerinnen und Schülern nachweisen, welche unterschiedliche Stärken und Schwächen in den verschiedenen Dimensionen des Wechsels von Darstellungsart und Repräsentationsform aufweisen. In der 8. Klasse hingegen erweis sich eine Modellierung mit Fähigkeitsklassen als empirisch tragfähig, da sich eine bedeutsame Interaktion von Clusterzugehörigkeit und Ausprägungen auf den Merkmalsdimensionen zeigt (s. Tabelle 8). Das 6-Cluster-Modell weist den geringsten CAIC auf. Auch die im Vergleich zu den übrigen Modellen niedrige Fehlklassifikationsrate weist dieses Modell als optimal aus.

Tabelle 8: Latente Klassenanalyse

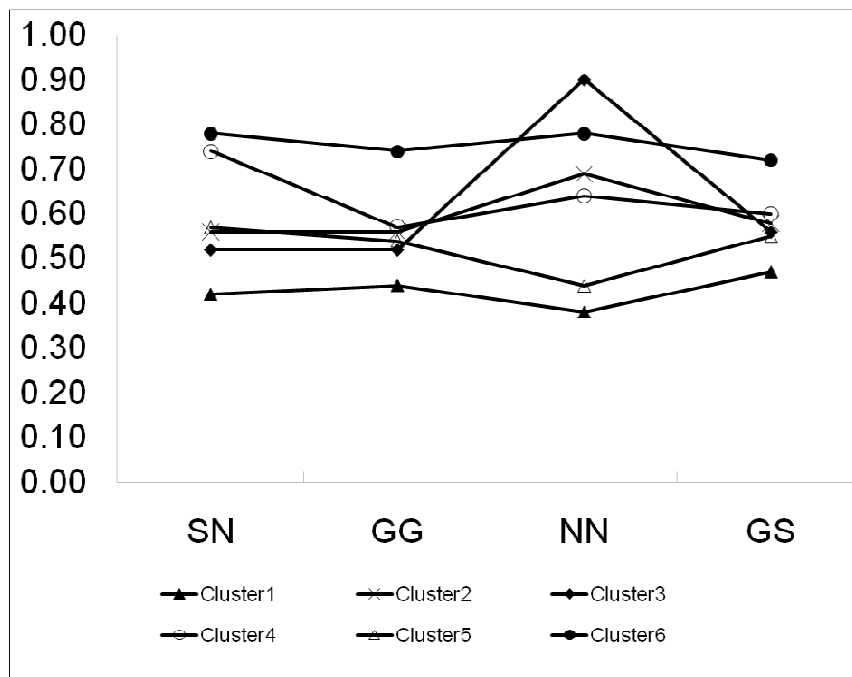
<i>Modell</i>	<i>LL</i>	<i>CAIC</i>	<i>BIC</i>	<i>Class. Err.</i>
1 Cluster	-3037.31	6131	6123	0.0000
2 Cluster	-2949.89	6021	6004	0.1418
3 Cluster	-2896.31	5978	5952	0.1622
4 Cluster	-2862.49	5975	5940	0.1828
5 Cluster	-2840.95	5996	5952	0.2138
6 Cluster	-2772.91	5924	5871	0.1189
7 Cluster	-2758.66	5960	5898	0.2083

Anmerkung: LL=Likelihood; CAIC= Consistent Akaike Information Criterion ;

BIC= Bayesian Information Criterion; Class. Err.= Fehlklassifikationsrate

Es zeigen sich typische Schülerprofile, die grafische Darstellung der Ergebnisse für die 8. Klasse findet sich in der folgenden Abbildung 13.

Abb.13: Kompetenzprofile in der 8. Klassenstufe



Anmerkung: SN=situativ-numerisch, NN=numerisch-numerisch, GG=grafisch-grafisch, SG=situativ-grafisch.

Die Cluster 1 (29 % der untersuchten Achtklässler/innen) und 6 (6 %) unterscheiden sich lediglich hinsichtlich des Grundniveaus über alle Dimensionen hinweg in ähnlicher Weise. Hier bildet sich also ein generell unterschiedliches Fähigkeitsniveau in allen vier Dimensionen in ähnlicher Weise ab. Für die übrigen Cluster zeigen sich jedoch auffällige typologische Strukturen, die unterschiedliche diagnostisch relevante Kompetenzstrukturen widerspiegeln. So repräsentiert beispielsweise Cluster 4 (15 %) Schülerinnen und Schüler, die bei ansonsten eher durchschnittlicher Leistung über eine hohe Kompetenz beim Repräsentationswechsel von der Situation ins Numerische (SN) sowie bei der Verarbeitung innerhalb der numerischen Repräsentation (NN) verfügen. Cluster 3 (15 %) hingegen ist

gekennzeichnet durch eine markante Stärke im Bereich der numerischen Verarbeitung (NN) bei sonst durchschnittlichen Leistungen. Schülerinnen und Schüler, die dem Cluster 5 (11 %) angehören, zeigen einen ähnlichen Profilverlauf, jedoch ist die Fähigkeit zur Verarbeitung innerhalb der numerischen Repräsentation nicht so hoch ausgeprägt wie in Cluster 3. Cluster 2 (21 %) ist gekennzeichnet durch eine relative Schwäche innerhalb der numerischen Verarbeitung. Es wird zu prüfen sein, ob diese Profile sich als stabil gegenüber der Wahl anderer Populationen oder gegenüber zeitlichen Entwicklungen erweisen.

4.4.2 Zusammenhang von Kompetenzprofilen und figuralen Denken

Zum besseren Verständnis dieser Kompetenzprofile wurde der Zusammenhang zwischen der Clusterzugehörigkeit und der Fähigkeit im figuralen Denken analysiert. Es zeigte sich, dass sich die verschiedenen Cluster signifikant ($F_{5, 385}=7,75$; $p<0,001$) und mit hoher Effektstärke ($\eta^2=0,09$) in ihren Mittelwerten im Untertest N2 im KFT 4-12+ R unterscheiden. Cluster 1 ($T=52,2$) und Cluster 3 ($T=51,9$) weisen beide vergleichsweise niedrige Werte figuralen Denkens auf (vgl. Tabelle 9), Cluster 3 ist jedoch durch eine hohe Kompetenz im Lösen von numerischen Aufgaben gekennzeichnet. Bei den grafischen Aufgaben und bei Aufgaben, die einen Repräsentationswechsel von der Situation ins Numerische verlangen, zeigen sich in diesem Cluster Schwächen. Der höchste Mittelwert im figuralen Denken zeigt sich in Cluster 6 ($T=60,7$), dieser überdurchschnittliche Wert spiegelt sich ebenfalls bei den hohen Kompetenzen in den verschiedenen Dimensionen des Mathematiktests wieder. Obwohl Cluster 4 relative Schwächen beim Bearbeiten von grafischen Aufgaben hat, ist der Wert im figuralen Denken ($T=57,3$) relativ hoch ausgeprägt.

Tabelle 9: Mittelwert der Cluster im Untertest N2 aus dem KFT 4-12+ R

Cluster	Mittelwert	Standardabweichung	N
1	52,1	7,6	90
2	56,8	9,1	108
3	51,8	9,2	64
4	57,8	6,3	73
5	57,2	9,6	40
6	60,7	5,0	10

Anmerkung: Die Mittelwerte sind in T-Werten angegeben

Diese Befunde hinsichtlich des Zusammenhangs von Clusterzugehörigkeit und figuralen Denken geben erste Hinweise auf den Einfluss von Moderatorvariablen auf unterschiedliche Kompetenzprofile von Schülerinnen und Schülern.

4.5 Diskussion und Ausblick

Die Ergebnisse zeigen, dass der Repräsentationstyp für den mathematischen Inhaltsbereich „funktionale Veränderung“ bedeutsam für die Kompetenzstruktur ist.

Auf der Basis eines vierdimensionalen Kompetenzstrukturmodells konnten in der 8. Klassenstufe Kompetenzprofile ermittelt werden, welche die grafische und numerische Modellierung von Situationen bzw. den Wechsel zwischen diesen Repräsentationen beinhalten. Es konnten Cluster von Schülerinnen und Schülern identifiziert werden, die typische Kompetenzprofile aufweisen. So zeigt beispielsweise eine Gruppe eine hohe Kompetenz beim Repräsentationswechsel von der situativen Darstellung in die numerische Repräsentation sowie innerhalb der numerischen Darstellung, während die Leistungen bei grafischen Aufgaben eher durchschnittlich sind. Bei einer anderen Gruppe mit einem niedrigen Wert im figuralen Denken wurde eine markante Stärke in der numerischen Verarbeitung festgestellt, jedoch Schwächen bei der Bearbeitung von grafischen Aufgaben.

Die Ergebnisse zum Zusammenhang zwischen den Kompetenzprofilen und figuralem Denken sind inkonsistent. Die Erwartung, dass hohe Kompetenzen bei der Bearbeitung von Aufgaben mit grafischem Inhalt mit einer hohen Leistung im figuralen Denken einhergehen, konnte nicht bestätigt werden. Möglicherweise sind die Befunde durch unterschiedliche Kompetenzniveaus der Schülerinnen und Schüler in der Bild-Text-Integration (vgl. Schnotz, 2005) zu erklären.

Die Ergebnisse von Snow und Yalow (1982) hinsichtlich der Präferenz schwächerer Schülerinnen und Schüler von grafischen präsentierten Aufgaben konnten nicht auf den Inhaltsbereich „funktionale Veränderung“ übertragen werden. Schwächere Schülerinnen und Schüler weisen keine Stärken bei grafisch repräsentierten Aufgaben auf.

In der 7. Klasse konnten keine typischen Profile von Schülerinnen und Schülern identifiziert werden, welche unterschiedliche Stärken und Schwächen hinsichtlich des Wechsels von Darstellungsart und Repräsentationsform aufweisen. Diese Schülerinnen und Schüler unterscheiden sich lediglich im Niveau ihrer Lösungshäufigkeit in den untersuchten Dimensionen.

Diese Befunde sind den Ergebnissen von Kleine (2005) ähnlich, der parallele Profilverläufe dargestellt hat. Möglicherweise handelt es sich hier um eine Polytomisierung eines eigentlich kontinuierlichen Fähigkeitsspektrums. Allerdings könnte dies auch darauf zurückzuführen sein, dass die Schülerinnen und Schüler noch am Anfang des Kompetenzerwerbs in diesem Bereich stehen und ihre Fähigkeiten noch nicht in dem Maß ausdifferenziert sind wie in der 8 Klasse. Es kann aber auch als Indiz für eine relative Homogenität des Curriculums zum Einstieg in diesen Bereich gewertet werden.

4.6 Grenzen der vorliegenden Studie und Ausblick

Insgesamt geben die Ergebnisse erste Hinweise auf das Vorliegen typischer Kompetenzprofile von Schüler/innen beim Umgang mit Funktionen. Der praktische Nutzen der identifizierten Profile hinsichtlich Diagnose und Fördermöglichkeiten auf Einzelschülerebene ist beim derzeitigen Stand der Forschung als noch nicht ausreichend zu bewerten. Die Analyse der Binnenstruktur des Kompetenzmodells zeigt noch unbefriedigende Reliabilitäten für drei der vier Dimensionen. Es besteht für diese durch eine zu geringe Anzahl von Items repräsentierten Konstrukte der Bedarf nach zusätzlichen und konstruktvalideren Items, um die psychometrischen Eigenschaften der Skalen zu optimieren. Auch für die Skala „situativ-grafisch“ (SG) mit akzeptabler Reliabilität müssen die Items in

Bezug auf fokussiertere Iteminhalte optimiert werden, um deren Trennschärfe sowie die Skalensvalidität und -reliabilität zu erhöhen. Auf der Basis dieses hinsichtlich psychometrischer Eigenschaften verbesserten Testinstruments lassen sich möglicherweise Profile von Schüler/innen identifizieren, die sich stärker in ihren Lösungshäufigkeiten der einzelnen Skalen unterscheiden. Diese Erkenntnisse könnten in einem Förderprogramm berücksichtigt werden, das im Rahmen einer formativen Diagnostik eingesetzt werden kann.

Zu den grundsätzlichen Zusammenhängen von spezifischen Kompetenzprofilen beim Repräsentationswechsel und einzelnen Moderatorvariablen haben sich noch nicht genügend Hinweise ergeben. Da sich keine Zusammenhänge der Leistung der Schüler/innen im Mathematiktest und möglichen emotionalen, bzw. motivationalen Variablen ergeben haben (vgl. 5.1), sollten weitere kognitive Variablen untersucht werden. Dies sollen z.B. verbale Fähigkeiten (vgl. Heller & Perleth, 2000) zur Überprüfung der Bedeutung des Aufgabenkontexts („Situation“) und weitere grafisch-räumliche Fähigkeiten sein (vgl. Jäger, Süß & Beauducel, 1997).

5 ZUSÄTZLICHE AUSWERTUNGEN

5.1 Analyse der Itemschwierigkeiten unter Einbeziehung von Moderatorvariablen anhand von „Differential Item Functioning“

Neben der Strukturanalyse der Kompetenz im mathematischen Inhaltsbereich „Funktionale Veränderung“ wurden explorativ verschiedene mögliche Moderatoren auf die Kompetenz und ihre Komponenten untersucht. Es wurden verschiedene Skalen, die zumeist aus Studien wie PISA oder BIQUA stammen, verwendet. Die Skalen und ihre Kennwerte finden sich im Anhang unter Abschnitt 9.7.

Der Zusammenhang zwischen den potentiellen Moderatorvariablen und der Leistung im Mathematiktest wurde im ersten Schritt durch eine Korrelationsanalyse ermittelt. Dazu wurden die einzelnen Skalen (z.B. Angst, Motivation, Selbstwirksamkeit, Note in Mathematik, Allgemeine Intelligenz) mit Leistungen der Schüler/innen im Mathematiktest korreliert.

Es zeigte sich, dass die Note in Mathematik signifikant mit der Leistung in den verschiedenen Bereichen des Tests zusammenhängt (vgl. Tabelle 10).

Die Skala „Figurales Denken“ (KFT) korreliert ebenfalls mit der Leistung der Schülerinnen und Schüler im Mathematiktest (siehe Tabelle 10). Herauszuheben ist dabei die erwartungskonforme, tendenziell höhere Korrelation der grafischen Dimensionen mit der Skala des figuralen Denkens“.

Tabelle 10: Korrelationen zwischen Moderatorvariablen

	Situativ- Numerisch	Grafisch- Grafisch	Numerisch- Numerisch	Grafisch- Situativ
Note	0.303(**)	0.274(**)	0.234(**)	0.342(**)
KFT- Wert	0.195(**)	0.253(**)	0.172(**)	0.280(**)

* Die Korrelation ist auf dem Niveau von $p=0,05$ (2-seitig) signifikant.

** Die Korrelation ist auf dem Niveau von $p=0,01$ (2-seitig) signifikant.

Für die erhobenen emotionalen und motivationalen Moderatorvariablen ließen sich keine Zusammenhänge mit der Mathematikleistung nachweisen.

Die vorgestellten Moderatorvariablen wurden anhand von *Differential Item Functioning* (DIF) analysiert. Unter DIF versteht man, dass bei getrennter Kalibrierung für verschiedene Teilgruppen (z.B. für Jungen und Mädchen) unterschiedliche Schwierigkeitsparameter resultieren. DIF liegt vor, wenn Personen mit gleicher latenter Merkmalsausprägung in einem Item unterschiedliche Lösungshäufigkeiten aufweisen. Die DIF-Analyse der Items wurde mit der Software ConQuest durchgeführt und es wurde von dem vierdimensionalen Modell (Repräsentationswechsel) ausgegangen. Als Kriterium für bedeutsames DIF wird ein Unterschied in den Itemparametern für die verglichenen Teilgruppen von mindestens 0,5 logits (Draba, 1977; Adams & Carstensen, 2002) zugrunde gelegt. Dieses Kriterium wird auch in der PISA Studie verwendet.

Ergebnisse für die Variable Geschlecht

In der PISA Studie 2000 wurden von Adams und Carstensen (2002) Geschlechtsunterschiede auf Gesamtskalenebene untersucht. Als bedeutsame Differenz zwischen den Lösungshäufigkeiten wird ein Unterschied von mehr als 0,5 logits angenommen. Allerdings wird hier den Unterschieden in den Einzelitems nicht inhaltlich auf Itemebene nachgegangen. Rost et al. (2003) analysierten die PISA-Daten auf Itemebene. Es ergab sich, dass Items, die leichter von Mädchen gelöst werden konnten, oftmals komplexe Bewertungs- und Entscheidungsprozesse beinhalten oder durch Aufgabenstellungen mit verbalen Inhalten charakterisiert sind. Jungen lösen Aufgaben leichter, in denen Physikkenntnisse von Vorteil sind oder auch räumliches Vorstellungsvermögen angesprochen wird. Auch scheint der Kontext bei diesen Aufgaben eher die Jungen angesprochen zu haben (z.B. Autos). Unsere Analyse kommt zu ähnlichen Ergebnissen hinsichtlich des Geschlechts: Jungen lösen Aufgaben leichter, in denen grafisches Verständnis gefordert ist, sowie Aufgaben, die kontexthaltig sind. Mädchen lösen Aufgaben leichter, die keinen Kontext beinhalten und verbal präsentiert wurden.

In der folgenden Tabelle finden sind die Items, bei denen sich nach geschlechtsspezifischer Skalierung die Itemparameter zwischen Jungen und Mädchen um mindestens 0,5 logits (Draba, 1977) unterscheiden.

Tabelle 11: DIF Analyse, Geschlecht

Leichter gelöst von Jungen	Leichter gelöst von Mädchen
it_21bGGA	it_14bNNA
it_35dSGB	it_24bSGB
it_71cGNB	it_150bSxGN
it_36bSGA	it_2bGSB
it_20aGGA	it_26bSNB

it_11GGCa	it_107SGB
it_21aGGA	it_158Sxgn
it_27cSGA	it_13bSGC
it_31aNNC	it_108GSB
it_100GSA	it_150aSxGN

Figurales Denken

Zur explorativen Analyse des Einflusses des figuralen Denkens wurde der Untertest „Figurales Denken“ des KFT eingesetzt, in dem Beziehungen zwischen grafischen Abbildungen hergestellt werden sollen. Zur Untersuchung des Einflusses dieser Fähigkeit auf die Schwierigkeit der Items wurden anhand eines Median-Splits zwei Gruppen gebildet. Der Median beträgt in der untersuchten Stichprobe $T=54$. Es zeigt sich, dass Schülerinnen und Schülern mit einem hohen Wert ($T>54$) Items leichter fielen als Schüler/innen mit niedrigerem Wert ($T<54$), in denen Grafen erzeugt oder interpretiert werden sollten. Die Items, die Kinder mit niedrigerem Wert im KFT leichter lösen, weisen kein einheitliches Muster auf.

Tabelle 12: DIF Analyse, figurales Denken

Leichter gelöst von Kindern mit $IQ<Median$	Leichter gelöst von Kindern mit $IQ>Median$
it_14bNNA	it_1aSGB
it_17GGB	it_20aGGA
it_34aSNBa	it_37dGSB
it_72aSGB	it_31aNNC
it_15GGB	it_37aSNB
it_72bSNB	
it_28bSGB	

Auch für die übrigen Skalen der Moderatorvariablen, wie Selbstreflektiertes Lernen, Angst oder Motivation wurden DIF-Analysen durchgeführt. Die beiden verglichenen Gruppen wurden anhand eines Median-Splits ermittelt. In diesen Analysen konnte kein einheitliches Muster ermittelt werden, das die jeweiligen Gruppen kennzeichnet.

Insgesamt konnte die DIF-Analyse lediglich für das Geschlecht und für das figurale Denken zur differenzierten Analyse der Kompetenzstruktur beitragen. Schüler/innen mit einem hohen Wert ($T > 54$) lösten solche Items leichter als Schüler/innen mit niedrigerem Wert ($T < 54$), in denen Grafen erzeugt oder interpretiert werden sollten. Möglicherweise spielt das räumliche Denken in beiden Fällen eine Rolle. Bei Items, die Kinder mit niedrigerem Wert im KFT leichter lösen, ergab sich kein einheitliches Muster. Jungen lösen kontexthaltige Aufgaben leichter und solche Items, in denen grafisches Verständnis gefordert ist. Mädchen lösen Aufgaben leichter, die wenig Kontext beinhalten und verbal präsentiert wurden (vgl. Rost et al., 2003). Die Ergebnisse sollten jedoch vorsichtig interpretiert werden, da unterschiedliche Rater der Aufgaben zu anderen Ergebnissen kommen könnten.

Für die erhobenen emotionalen und motivationalen Moderatorvariablen, wie z.B. für das selbstreflektierte Lernen konnten keine korrelativen Zusammenhänge mit der Mathematiknote ermittelt werden. Dieses Ergebnis deckt sich mit den Resultaten einer anderen Studie, in der der Zusammenhang von Mathematikleistung und selbstreflektiertem Lernen untersucht wurde (Reiss, 2009). Auch hier wurde kein Zusammenhang zwischen Mathematikleistung und selbstreflektiertem Lernen gefunden.

Der Einfluss von emotionalen und motivationalen Variablen auf die Mathematikleistung erscheint also auf Gruppenebene als gering. Zur differenzierten Analyse der Kompetenzstruktur beim Umgang mit Funktionen wäre zu empfehlen, zukünftig verstärkt

den Einfluss kognitiver Variablen, wie z.B. verbale oder numerische Fähigkeiten auf die Leistungen der Schüler/innen in Mathematik zu untersuchen.

5.2 Empirische Bestimmung von Kompetenzniveaus in der Dimension

„Wechsel von Situation und grafischer Repräsentation“ (SG)

Zur weiteren Ausdifferenzierung des in Kapitel 2 vorgeschlagenen Kompetenzmodells soll die Dimension „Wechsel Situation-grafisch“ empirisch in Kompetenzniveaus eingeteilt werden. Dadurch ist es möglich, beispielweise eine inhaltliche Rückmeldung über die erbrachten Leistungen zu geben oder längsschnittlich Kompetenzentwicklungen abzubilden.

In Studien wie TIMSS und etwas abgewandelt auch in PISA wurde die Methode des „Scale anchoring“ (Beaton & Allen, 1992) verwendet. Durch eine nachträgliche Analyse der Aufgaben anhand ihrer Schwierigkeiten wurden eindimensionale Skalen intuitiv in Kompetenzniveaus unterteilt (OECD, 2005). Die Zahl und die Breite der Kompetenzniveaus wurden dabei meist willkürlich festgelegt. Dieses Verfahren ist als unbefriedigend zu bewerten, daher wird hier eine Methode verwendet, die Hartig (2007) im Rahmen der DESI-Studie entwickelt hat. Anzahl und Breite der Kompetenzniveaus werden in diesem Ansatz nicht willkürlich festgelegt, sondern nach empirischen Kriterien bestimmt. Für die Definition von Niveaus sind zwei separate Analyseschritte notwendig: Zunächst werden die empirischen Aufgabenschwierigkeiten anhand einer Raschanalyse ermittelt. Hier werden die Schwierigkeitsparameter des vierdimensionalen Kompetenzmodells (vgl. Kapitel 2) verwendet. Anschließend werden diese Schwierigkeiten mittels einer Regressionsanalyse anhand von Aufgabenmerkmalen vorhergesagt.

Dazu wurden a priori Aufgabenmerkmale entwickelt, welche bei der Schwellenfindung der Kompetenzskala dienlich sind. Zusätzlich charakterisieren diese Merkmale die Anforderungen und Schwierigkeiten des Tests. Zur Einschätzung der Aufgabenschwierigkeiten wurden vier wesentliche Merkmale identifiziert: der Kontext, der

Repräsentationswechsel (Ainsworth et al., 2002), die Verknüpfung von Informationen und der funktionale Aspekt (Vollrath, 1989; Malle, 2000).

Kontext (KO)

Der Kontext einer Aufgabe wird in zwei Stufen eingeteilt: *einfach* oder *komplex*. Bei einem einfachen Kontext handelt es sich um ein Thema aus der Lebenswelt der Schüler/innen, der leichter zu verstehen ist, als eine komplexe Situation, die nicht unmittelbar im Zentrum des Alltags der Schüler steht.

- 0 Einfacher, bekannter Kontext, z.B. Aus dem Alltag der Schüler/innen; einfache Einkleidung ohne komplexe Zusatzinformationen;
- 1 Kontext ist komplex, z.B. Bakterienwachstum, aspektreiche Realsituation, über die Wissen nötig ist jenseits von Alltagswissen.

Repräsentationswechsel (RW)

In der Skala „Wechsel Situation-grafisch“ sind nur Aufgaben enthalten, die einen Repräsentationswechsel von der Situation ins Grafische oder umgekehrt erfordern. Es wird bei der Klassifikation der Aufgabenmerkmale unterschieden, ob ein oder mehrere Wechsel zum Lösen eines Items erforderlich sind.

- 0 Repräsentationswechsel von der Situation ins Grafische oder umgekehrt;
- 1 Mehrere Wechsel zwischen Situation und Graf erforderlich.

Verknüpfung von Informationen (VI)

Das Merkmal „Verknüpfen von Informationen“ bezieht sich auf die Komplexität einer Aufgabe. Die erste Stufe umfasst die einfache Entnahme einzelner Werte aus der Aufgabe, für die zweite ist es notwendig mehrere Informationen zu verknüpfen.

- 0 Einfache Entnahme von Information, z.B. einzelne Werte aus Grafen oder Tabellen entnehmen;
- 1 Verknüpfung von mehreren Informationen erforderlich, z.B. begründen, warum Zusammenhang proportional ist;

Funktionaler Aspekt (FA)

Zur Einschätzung der Schwierigkeit eines Items wurden schließlich die wesentlichen Aspekte von Funktionen nach Vollrath (1989) herangezogen (vgl. Kapitel 1.2.2). Die erste Klasse enthält nur den Einsetzungsaspekt während bei der zweiten Klasse der Kovarianzaspekt eine Rolle spielt oder auch die Funktion als Ganzes analysiert werden muss.

- 0 Nur der Einsetzungsaspekt ist notwendig;
- 1 Der Kovariationsaspekt ist erforderlich (zwei Variablen) oder die Funktion ist als Objekt analysieren.

Zur Einteilung der Skala „Wechsel Situation-Grafisch“ in Kompetenzniveaus wurden die Merkmale Kontext, Repräsentationswechsel, Informationsverknüpfung und die funktionalen Aspekte mittels einer Regressionsanalyse überprüft (vgl. Hartig 2007). In der Tabelle 13 sind die Regressionsgewichte der Merkmale bei der Vorhersage der

Schwierigkeiten der Items aufgeführt. Mit $R^2=.55$ erklärt das Regressionsmodell etwas mehr als die Hälfte der beobachtbaren Unterschiede zwischen den Aufgabenschwierigkeiten, was sich mit den Ergebnissen anderer Untersuchungen deckt (vgl. Harsch & Schröder, 2007).

Tabelle 13: Regressionsgewichte bei der Vorhersage der Aufgabenschwierigkeit ($R^2=.553$)

Merkmale	Regressionsgewicht	Standardfehler	Irrtumswahrscheinlichkeit
Regressionskonstante	-1.591	.358	<.000
Kontext	.725	.409	<.088
Repräsentationswechsel	.546	.409	<.194
Informationsverknüpfung	.910	.449	<.054
Funktionaler Aspekt	.285	.578	<.626

Um Schwellen zwischen den Kompetenzniveaus abzuleiten, werden die aus dem Regressionsmodell abgeleiteten Schwierigkeiten bestimmter Kombinationen von Aufgabenmerkmalen benutzt (vgl. Hartig, 2007; Harsch & Schröder, 2007). Die Kompetenzskala „Wechsel Situation-Grafisch“ besteht aus fünf Kompetenzniveaus, die mit A bis D bezeichnet sind. Die Leistungen der Schüler/innen können einerseits gemäß ihrer Fähigkeiten auf einem bestimmten Niveau eingeordnet werden, andererseits ist es möglich, die Items hinsichtlich ihrer Schwierigkeiten und Merkmale einzuordnen.

In der Tabelle 14 finden sich die die Schwellenwerte der Kompetenzniveaus und die jeweiligen Merkmalskombinationen der vier ins Regressionsmodell eingegangenen Aufgabenmerkmale.

Tabelle14: Schwellenwerte und Merkmalskombinationen der Kompetenzniveaus der Skala SG

Kompetenz-niveau	Erwartete Aufgabenschwierigkeit in Logits	Kontext	Repräsentationswechsel	Informationsverarbeitung	Funktionaler Aspekt
Unter A	- 1, 60	0	0	0	0
A	-1, 30	0	0	0	1
B	-0, 39	0	0	1	1
C	0, 15	0	1	1	1
D	0, 88	1	1	1	1

Die empirisch ermittelten Kompetenzniveaus werden im Folgenden anhand ihrer Merkmale inhaltlich beschrieben:

Kompetenzniveau A:

Schüler/innen auf Kompetenzniveau A können Aufgaben lösen, bei denen nur der Einsetzungsaspekt von Funktionen von Bedeutung ist. Der Kontext der Aufgaben ist bekannt, und es sind keine komplexen Verknüpfungsleistungen, bzw. Repräsentationswechsel erforderlich.

Kompetenzniveau B:

Von Personen auf Kompetenzniveau B können Aufgaben dann gelöst werden, für die die Verknüpfung von mehreren Informationen notwendig ist und welche den Einsetzungsaspekt von Funktionen enthalten.

Kompetenzniveau C:

Schülerinnen und Schüler auf Kompetenzniveau C beherrschen mehrere Repräsentationswechsel bei der Bearbeitung von funktionalen Zusammenhängen und können mehrere Informationen verknüpfen. Zusätzlich wird der Kovarianzaspekt angewendet oder auch die Funktion als Ganzes analysiert.

Kompetenzniveau D:

Das höchste Kompetenzniveau D ist charakterisiert durch Aufgaben, die zusätzlich zu den Merkmalen in Niveau C einen komplexen Kontext beinhalten.

Diskussion

Bei der empirischen Bestimmung von Kompetenzniveaus fällt auf, dass vorwiegend der untere Durchschnittsbereich differenziert erfasst wird. Da es sich bei der Stichprobe um Gymnasiasten handelt, werden von vielen Schüler/innen alle postulierten schwierigkeitsbestimmenden Merkmale beherrscht.

Diese erste Einteilung der Dimension „Wechsel Situation-Grafisch“ in Kompetenzniveaus sollte in folgenden Studien überprüft werden. Dazu sollten die postulierten Aufgabenmerkmale empirisch validiert und um schwierigere Merkmale ergänzt werden. Zusätzlich sollten mehrere Rater eingesetzt werden, welche die Aufgaben hinsichtlich der Schwierigkeitsmerkmale einschätzen und die Beurteilerübereinstimmung ermittelt werden (vgl. Wirtz & Caspar, 2002). Es sollte auch anhand eines Expertenrankings eingeschätzt werden, wie gut die Aufgaben inhaltlich in den jeweiligen Niveaustufen passen, d.h. inwieweit die schwierigkeitsbestimmenden Faktoren auf die Items eines Niveaus zutreffen.

Zur Beurteilung der hier verwendeten Methode ist anzumerken, dass es sich bei dem linear-additiven Modell um ein relativ einfaches Vorhersagemodell handelt. Alternative Modelle zu Vorhersage der Schwierigkeiten der Items sind denkbar, etwa das Linear-Logistische Testmodell (LLTM; z.B. Fischer, 1996). Im LLTM wird eine gewichtete Kombination von Basisparametern zur Erklärung der Aufgabenschwierigkeit eingesetzt, wobei jedes Aufgabenmerkmal durch einen Basisparameter repräsentiert wird. Diese könnten analog zu den in der vorliegenden Studie gewählten Regressionsgewichten bestimmt werden. Hartig und Frey (2005) verglichen die Ergebnisse zur Bildung von Kompetenzniveaus in der DESI-Studie auf der Basis einer Regressionsanalyse sowie anhand eines LLTMs und kamen zu ähnlichen Ergebnissen.

Aus Gründen der Nachvollziehbarkeit der Ergebnisse wurde daher hier die Methode der separaten Raschskalierung und Regressionsanalyse gewählt.

6 GESAMTDISKUSSION

6.1 Modellierung mathematischer Kompetenz- ein Fazit

Die Ergebnisse der verschiedenen Studien wurden jeweils am Ende des betreffenden Kapitels diskutiert und werden im Folgenden vernetzt diskutiert. Die Zielsetzung der vorliegenden Arbeit war eine differenzierte Analyse der Kompetenzstruktur von Schüler/innen beim Problemlösen mit verschiedenen Repräsentationen von Funktionen. Dazu wurde ein Kompetenzstrukturmodell für den Mathematikunterricht entwickelt und empirisch überprüft. Zur differenzierten Analyse der Kompetenzstruktur wurden Abweichungen systematisch analysiert und Typen von Schüler/innen mit bestimmten Leistungsprofilen postuliert. Dadurch soll die Basis für ein formatives diagnostisches Instrument geschaffen werden, das im Mathematikunterricht eingesetzt werden kann.

6.1.1 Dimensionale Analyse der Kompetenzstruktur

Im Rahmen von „large-scale-assessment“-Studien, wie TIMSS oder PISA (z.B. OECD, 2005) wurden bisher eindimensionale Kompetenzmodelle zur vergleichenden Leistungsmessung von Gruppen auf Länderebene entwickelt, welche in diesem Kontext nützlich sind (vgl. Kapitel 1.1.3). Anhand einer dimensional Analyse der Kompetenzstruktur von Schüler/innen beim Umgang mit verschiedenen Repräsentationen von Funktionen wurde auf der Basis von modernen psychometrischen Methoden ein mehrdimensionales Kompetenzstrukturmodell vorgeschlagen und empirisch validiert (vgl. Kapitel 2). Kritisch anzumerken ist, dass die Reliabilität von drei der vier Dimensionen als noch nicht ausreichend beurteilt werden muss. Diese liegt im Bereich von 0,52 bis 0,64 und sollte eigentlich über 0,7 liegen (vgl. Rauch & Hartig, 2007). Teilweise ist dies durch eine

geringe Zahl von Items zu erklären, welche die einzelnen Dimensionen repräsentieren. Kaiser und Wilson (2004) empfehlen mindestens 9 Items pro Dimension, bei einer Skala des eingesetzten Testinstruments lag die Zahl der Items nur bei fünf (vgl. Kapitel 2.2). Die Konstruktion von Kompetenzstrukturmodellen ist eine neue Forschungsrichtung in der empirischen Bildungsforschung. Die geringe Reliabilität der Skalen kann daher möglicherweise auch darauf zurückzuführen sein, dass noch nicht genügend Erfahrung mit der Konstruktion reliabler Skalen in der differenzierten Erfassung eines fokussierten Kompetenzbereichs besteht.

Die latenten Korrelationen der vier Dimensionen des Kompetenzstrukturmodells sind erwartungsgemäß hoch (0,57-0,76), aber noch akzeptabel. In der PISA-Studie wurden höhere latente Korrelationen zwischen verschiedenen Domänen gemessen (Carstensen, 2008).

Der Itempool sollte bezüglich der identifizierten Dimensionen optimiert werden, um eine höhere Reliabilität der einzelnen Skalen zu gewährleisten. Wie die zweite Studie (vgl. Kapitel 3) zur weiteren dimensional Analyse der Kompetenzstruktur gezeigt hat, ist ein Modell mit Mehrfachladungen („Within-Item-Multidimensionality“) dem in Kapitel 2 vorgeschlagenen Kompetenzmodell nicht überlegen. Diese Studie macht deutlich, wie schwer kognitive Prozesse bei Items zu modellieren sind, die anhand verschiedener Repräsentationsformen, wie z.B. der tabellarischen oder grafischen Repräsentation gelöst werden können. Bei einer Optimierung des Itempools sollten daher verstärkt fokussierte Items konstruiert werden, die nur jeweils eine der vier Dimensionen des Kompetenzmodells abbilden. Zusätzlich sollten die Erkenntnisse, die aus der Einteilung der Skala „Wechsel von Situation und grafischer Repräsentation“ hinsichtlich der schwierigkeitsbestimmenden Merkmale gewonnen wurden, berücksichtigt werden (vgl. 5.1). Bei der empirischen Bestimmung von Kompetenzniveaus wurde vorwiegend der untere Durchschnittsbereich

differenziert erfasst. Da es sich bei der Stichprobe um Gymnasiasten handelt, wurden die postulierten schwierigkeitsbestimmenden Merkmale von vielen Schüler/innen beherrscht. Daher sollten basierend auf Erkenntnissen der Literatur zusätzlich Merkmale erhoben werden, die für schwierige Items charakteristisch sind.

6.1.2 Analyse der Kompetenzstruktur anhand von Kompetenzprofilen und Abweichungen von der dimensionalen Struktur

Zusätzlich zur dimensionalen Analyse der Kompetenzstruktur beim Repräsentationswechsel mit Funktionen wurden weitere differenzierte Strukturanalysen vorgenommen. In Studie 3 (vgl. Kapitel 4) wurden auf der Basis des in Kapitel 2 entwickelten Kompetenzstrukturmodells Kompetenzprofile von Schülerinnen und Schülern postuliert. Die Ergebnisse geben erste Hinweise auf das Vorliegen typischer Kompetenzprofile von Schüler/innen beim Umgang mit Funktionen. Der praktische Nutzen der identifizierten Profile hinsichtlich Diagnose und Fördermöglichkeiten auf Einzelschülerebene ist beim derzeitigen Stand der Forschung als noch nicht ausreichend zu bewerten. Auf der Basis eines hinsichtlich psychometrischer Eigenschaften verbesserten Testinstruments lassen sich möglicherweise Cluster von Schüler/innen identifizieren, die sich stärker in ihren Lösungshäufigkeiten der einzelnen Skalen unterscheiden.

Zu den grundsätzlichen Zusammenhängen von mathematischer Kompetenz beim Repräsentationswechsel bei funktionalen Zusammenhängen und einzelnen Moderatorvariablen, wie Motivation, Angst oder selbstreflektiertes Lernen haben sich noch nicht genügend Hinweise ergeben (vgl. Kapitel 5.1). Für die erhobenen emotionalen und

motivationalen Moderatorvariablen konnten keine korrelativen Zusammenhänge mit der Mathematiknote ermittelt werden (vgl. Reiss, 2009).

Die vorgestellten Moderatorvariablen wurden auch anhand von Differential Item Functioning (DIF) analysiert. DIF liegt vor, wenn Personen mit gleicher latenter Merkmalsausprägung in einem Item unterschiedliche Lösungshäufigkeiten aufweisen. Die DIF-Analyse konnte lediglich für das Geschlecht und für das figurale Denken zur differenzierten Analyse der Kompetenzstruktur beitragen. Schüler/innen mit einem hohen Wert ($T > 54$) fielen Items leichter, in denen Grafen erzeugt oder interpretiert werden sollten als Schüler/innen mit niedrigerem Wert ($T < 54$). Möglicherweise spielt das räumliche Denken sowohl beim Lösen der figuralen Items des KFT als auch bei der Bearbeitung von Aufgaben mit grafischem Inhalt eine Rolle. Hinsichtlich des Geschlechts ergaben sich die folgenden Ergebnisse: Jungen lösen kontexthaltige Aufgaben leichter und solche Items, in denen grafisches Verständnis gefordert ist. Mädchen lösen Aufgaben leichter, die wenig Kontext beinhalten und verbal präsentiert wurden (vgl. Rost et al., 2003b).

Das Verfahren der DIF-Analyse scheint –zumindest bei kognitiven Variablen und beim Geschlecht- geeignet, um Kompetenzstrukturen jenseits der dimensional Struktur eines Kompetenzbereichs weiter aufzuklären. Der Zusammenhang von emotionalen sowie motivationalen Variablen und Mathematikleistung erscheint auf Gruppenebene insgesamt als gering (vgl. Reiss, 2009). Zur differenzierten Analyse der Kompetenzstruktur beim Umgang mit Funktionen wäre zu empfehlen, zukünftig verstärkt den Einfluss kognitiver Variablen zu untersuchen. Dies können z.B. verbale Fähigkeiten (vgl. Heller & Perleth, 2000) zur Überprüfung der Bedeutung des Aufgabenkontexts („Situation“) und weitere grafisch-räumliche Fähigkeiten sein (vgl. Jäger, Süß & Beauducel, 1997).

Insgesamt konnte gezeigt werden, dass eine mehrdimensionale Modellierung der Kompetenzstruktur beim Umgang mit verschiedenen Repräsentationen von Funktionen prinzipiell zu leisten ist. Auch durch weitere Analyse der Kompetenzstruktur anhand von Kompetenzprofilen und einer Analyse der Itemschwierigkeiten in unterschiedlichen Gruppen konnten erste Ergebnisse erzielt werden.

6.2 Implikationen für die Unterrichtspraxis

Es werden nachfolgend aus den zentralen Befunden der vorliegenden Arbeit einige Implikationen für die Diagnose und Förderung der mathematischen Kompetenz beim Problemlösen mit Funktionen abgeleitet.

Der Repräsentationswechsel von der grafischen bzw. numerischen Repräsentation und Situation, sowie die Verarbeitung innerhalb der grafischen und numerischen Darstellung wurden als differenzierte Dimensionen mathematischer Kompetenz beim Umgang mit Funktionen ausgewiesen (vgl. Swan, 1985; Ainsworth et al., 2002). Damit konnte empirisch belegt werden, dass die Fähigkeit, Text und Graf als verschiedene Repräsentationsformen zu verarbeiten und auf die jeweils andere Repräsentationsart zu beziehen (Schnotz, 2005) für den Umgang mit Funktionen relevant ist.

Grundsätzlich bedeutet es, dass diese einzelnen Dimensionen bei der Diagnose und Förderung nicht gleichgesetzt, sondern einzeln berücksichtigt werden sollten. Die Analyse der Relationen der einzelnen Dimensionen des Kompetenzstrukturmodells gibt andererseits Hinweise darauf, dass Zusammenhänge der verschiedenen Kompetenzdimensionen bestehen.

Anhand eines Diagnostikums, das auf Grundlage des empirisch abgesicherten Kompetenzstrukturmodells entwickelt werden kann, lassen sich gezielt Stärken und Schwächen der Schüler/innen beim Repräsentationswechsel mit Funktionen auf Individual- und Gruppenebene identifizieren. Die Reliabilitäten der einzelnen Skalen des Testinstruments sollten über 0,7 liegen (vgl. Rauch et al., 2007), damit man von reliablen Messungen der Kompetenz ausgehen kann.

Auf der Basis des entwickelten Kompetenzmodells konnten in der 8. Klassenstufe Cluster von Schülerinnen und Schülern identifiziert werden, die typische Kompetenzprofile

aufweisen. So zeigt beispielsweise eine Gruppe eine hohe Kompetenz beim Repräsentationswechsel von der situativen Darstellung in die numerische Repräsentation sowie innerhalb der numerischen Darstellung, während die Leistungen bei grafischen Aufgaben eher durchschnittlich sind. Bei einer anderen Gruppe wurde eine Stärke in der numerischen Verarbeitung festgestellt, jedoch Schwächen bei der Bearbeitung von grafischen Aufgaben (vgl. Kapitel 4.4). Daraus lassen sich Ansätze für eine systematische Förderung mathematischer Kompetenz entwickeln.

Es hat sich in empirischen Untersuchungen gezeigt, dass die unterschiedlichen Repräsentationsformen mathematischer Zusammenhänge und deren Wechsel entscheidend für individuelle Lernprozesse sind. Damit sich ein positiver Lerneffekt einstellt, müssen die Schülerinnen und Schüler in der Lage sein, zwischen den vorliegenden Repräsentationen zu wechseln, die verbindenden Elemente zu identifizieren und zu verknüpfen (Seufert, 2003; Ainsworth et al., 2002).

Diese Erkenntnisse sollten in einem Förderprogramm berücksichtigt werden. Beispielsweise könnte die Fähigkeit zwischen den Repräsentationsformen Graf bzw. Tabelle und Text zu wechseln, gezielt trainiert werden, indem Schüler/innen lernen, die verbindenden Elemente zu identifizieren. In einem weiteren Schritt könnte vermittelt werden, die verbindenden Elemente von Text und Graf, bzw. Tabelle miteinander zu verknüpfen (vgl. Barzel et al., 2005; Herget et al., 2000). Ein weiterer Baustein des Förderprogramms könnte die Förderung der Fähigkeit der Verarbeitung innerhalb der grafischen und numerischen Darstellung sein.

Es empfiehlt sich in jedem Falle, den Wechsel der Repräsentation und die Verarbeitung innerhalb der grafischen und numerischen Repräsentation beim Problemlösen

mit Funktionen in den mathematischen Unterricht einzubetten und den Schüler/innen deren Relevanz zu verdeutlichen.

6.3 Schlussfolgerungen und Ausblick

Es ergeben sich aus den in der Diskussion abgeleiteten Implikationen für die Unterrichtspraxis sowie aus der Diskussion der einzelnen Studien verschiedene Ansätze für weiterführende Forschungsarbeiten auf dem Gebiet der Kompetenzmodellierung. Es konnte gezeigt werden, dass eine Modellierung der mathematischen Kompetenzstruktur, die sich auf das Problemlösen mit verschiedenen Repräsentationen von Funktionen bezieht, möglich ist

Wie bereits in der Diskussion angesprochen, wäre es die Aufgabe weiterer Forschung zunächst die Items zu optimieren, um eine höhere Reliabilität der einzelnen Dimensionen des Modells zu gewährleisten. Aus der Basis eines solchen hinsichtlich Reliabilität und Validität stabilen Modells lassen sich weitere Kompetenzstrukturanalysen durchführen und Erklärungsmodelle für unterschiedliche Kompetenzen und Kompetenzprofile bei Umgang mit Funktionen entwickeln (vgl. Harsch & Schröder, 2007; Rauch et al., 2007).

Zusätzlich können das optimierte Kompetenzstrukturmodell und die gefundenen Kompetenzprofile die Basis für ein Diagnoseinstrument mathematischer Problemlösefähigkeit beim Umgang mit Funktionen bilden. Anhand dieses Instruments lassen sich Erkenntnisse zu Förderbedarf und Förderungsmöglichkeiten auf Lerngruppen- und Individuenebene gewinnen. Das Diagnoseinstrument kann in späteren Studien dazu eingesetzt werden, um die Wirksamkeit von Fördermaßnahmen untersuchen. Es wäre sinnvoll, das Diagnoseinstrument aus ökonomischen Gründen computerbasiert einzusetzen.

Als weitere Perspektive ist die Modellierung von Längsschnittverläufen der Kompetenz von Lernenden beim Problemlösen mit Funktionen zu sehen (vgl. von Davier, 2005). Ein Ziel der aktuellen Bildungsforschung ist es, Kompetenzentwicklungsverläufe von Schülerinnen und Schülern aufzudecken und anhand von Kompetenzentwicklungsmodellen

zu beschreiben (Seidel & Prenzel, 2008). Dazu kann das Diagnoseinstrument den Ausgangspunkt liefern. Das Kompetenzstrukturmodell wurde in der vorliegenden Arbeit für die Klassenstufe 7 und 8 empirisch abgesichert. Es ist anzunehmen, dass die Darstellungsformen von Funktionen und der Wechsel zwischen verschiedenen Repräsentationsformen auch in höheren Klassenstufen eine entscheidende Rolle beim Problemlösen mit funktionalen Zusammenhängen spielt. Die Klärung der Frage, ob sich diese Strukturen auch in höheren Klassenstufen nachweisen lassen, wäre Aufgabe der weiteren Forschung. Das Modell könnte in diesem Fall auf die Klassenstufen 9 und 10 erweitert werden, indem beispielsweise der Umgang mit Termen als weitere Dimension angenommen wird.

Schließlich ergibt sich die Frage nach der Domänenspezifität der Kompetenz bei Umgang mit verschiedenen Darstellungsarten von Funktionen und dem Wechsel zwischen unterschiedlichen Repräsentationen. Auch im naturwissenschaftlichen Unterricht werden funktionale Darstellungen eingesetzt, etwa Geschwindigkeit-Zeit-Diagramme in der Physik. Für den Umgang mit solchen Funktionen erscheint das Verständnis dahinterliegende Prozesse ebenfalls grundlegend. Es stellt sich die Frage, ob sich einige Erkenntnisse zur dimensional Struktur mathematischer Kompetenz beim Umgang mit Funktionen, auf andere Fächer übertragen lassen. Zukünftige Forschungsarbeiten zur Kompetenzmodellierung im naturwissenschaftlichen Unterricht könnten das in der vorliegenden Arbeit beschriebene Kompetenzstrukturmodell weiterentwickeln und fächerübergreifend absichern.

7 ZUSAMMENFASSUNG

In der vorliegenden Arbeit wurde ein Kompetenzstrukturmodell im mathematischen Inhaltsbereich „funktionale Veränderung“ konstruiert, welches die Fähigkeiten von Schülerinnen und Schülern der Sekundarstufe I beim Problemlösen mit Darstellungsformen (numerisch, grafisch und situativ) und dem Wechsel zwischen ihnen beschreibt.

Zur empirischen Überprüfung des Modells wurde ein bereichsspezifischer Pool von 80 Items konstruiert, pilotiert und an einer Stichprobe von 872 Schülerinnen und Schülern der 7. und 8. Klasse aus Gymnasien in Baden-Württemberg und Hessen überprüft. Es wurde ein Multi-Matrix-Item-Sampling-Design verwendet, so dass jedem Schüler nur 34 Items anhand eines Paper-Pencil-Tests präsentiert wurden. Anhand ein- und mehrdimensionaler Rasch-Analysen („*between-item-multidimensionality*“) wurden verschiedene theoretisch abgeleitete Kompetenzstrukturmodelle daraufhin überprüft, ob sich die jeweils postulierten Strukturen in den Daten widerspiegeln. Durch Anwendung des „*multidimensional random coefficient multinomial logit*“ Modells (MRCML: Adams et al., 1997), das in der Software ConQuest implementiert ist, wurde ein empirischer Modellvergleich durchgeführt.

Ein vierdimensionales Modell, das die Wechsel zwischen und innerhalb von Repräsentationsformen (situativ-grafisch, situativ-numerisch, grafisch-grafisch, numerisch-numerisch) als wesentliche Kompetenzdimensionen annimmt, weist die beste Datenvorhersage auf. Die Separierbarkeit dieser Dimensionen wird weiterhin durch vergleichsweise moderate latente Interkorrelationen unterstützt ($r = .57$ bis $.76$). Jedoch konnte lediglich für den Repräsentationswechseltyp ‚situativ-grafisch‘ mit $.72$ eine akzeptable EAP/PV-Reliabilität erreicht werden. Trotz der validierten dimensional Struktur muss somit eine weitere Optimierung und Schärfung der Itemgruppen erfolgen, um den

psychometrisch fundierten Einsatz des Diagnostikums (insbesondere zur Individualdiagnostik) empfehlen zu können.

Zusätzlich wurde dieses Modell anhand weiterer Modelle („*within-item-multidimensionality*“) validiert, die Lösungswege der Schülerinnen und Schüler mit berücksichtigen. Auch die Ergebnisse dieses Modellvergleichs weisen das postulierte vierdimensionale Kompetenzstrukturmodell als das empirisch am besten passende Modell aus.

Zur Erfassung von intra-individuellen Unterschieden bei der Abbildung von Kompetenzentwicklungen wurde die Dimension „Repräsentationswechseltyp: situativ-grafisch“ exemplarisch in Kompetenzniveaus unterteilt. Nach einer Optimierung des Itempools, insbesondere, wenn eine zufriedenstellende Reliabilität der Skalen erreicht wurde, sollen die anderen Skalen in Kompetenzniveaus unterteilt werden.

Durch Analyse mittels *Differential Item Functioning* konnte gezeigt werden, dass die Schwierigkeit der Skalenitems in Einzelfällen systematisch mit den Ausprägungen im figuralen Denken variiert. Mittels latenter Klassenanalyse (Rost, 2003) konnte weiterhin nachgewiesen werden, dass man in der 8. Klassenstufe über die vier Dimensionen hinweg 6 typische Kompetenzprofile identifizieren kann. Diese Befunde zeigen, dass neben den dimensional Kompetenzaspekten auch typologische Strukturen durch den Itempool identifiziert werden können. Diese können Hinweise auf typische Kompetenzstrukturen geben, die im Sinne einer formativen Diagnostik zur Identifikation von spezifischem Förderbedarf oder Leistungsprofilen eingesetzt werden können.

Insgesamt weisen die Ergebnisse darauf hin, dass Wechsel von sprachlich-symbolhaften Darstellungen und grafisch-bildhaften Repräsentationen kognitive Strukturen

erklären können, die beim mathematischen Problemlösen eine Rolle spielen. Das postulierte Kompetenzmodell kann die Basis für ein Diagnoseinstrument mathematischer Problemlösefähigkeit beim Umgang mit Funktionen bilden und gleichzeitig Ansatzpunkte für die Förderung bieten. Zusätzlich wurden grundlegende methodische Vorgehensweisen zur curricular validen Erfassung von Kompetenzstrukturen anhand von multidimensionalen Modellen für die Bildungsforschung entwickelt. Damit konnte in einem curricular zentralen Kompetenzbereich der Grundstein für eine differenzierte Kompetenzmessung gelegt werden.

7 LITERATURVERZEICHNIS

- Adams, R. J. & Khoo, S. H. (1996). *Quest [Computer software]*. Melbourne, Australia: Australian Council for Educational Research.
- Adams, R. J., & Carstensen, C. H. (2002). Scaling Outcomes. In Adams, R. & Wu, M. (Eds.), *PISA 2000 Technical Report* (149-162). Paris: OECD.
- Adams, R. J., Wilson, M., Wang, W. (1997): The multidimensional random coefficient multinomial logit model. *Applied Psychological Measurement*, 21, 1-13.
- Ainsworth, S.E., Bibby, P.A & Wood, D.J. (2002). Examining the effects of different multiple representational systems in learning primary mathematics. *Journal of the Learning Sciences*. 11(1), 25-62.
- Baker, F. (2001). *The Basics of Item Response Theory* (2nd. Ed). University of Maryland, College Park, MD: ERIC Clearinghouse on Assessment and Evaluation.
- Barzel, B., Hußmann, S. & Leuders, T. (2005). Der "Funktionenführerschein". *Praxis der Mathematik in der Schule*. 47 (2), 20-25.
- Baumert, J., Evans, R. H. & Geiser, H. (1998). Technical problem solving among 10-year-old students as related to science achievement, out-of-school experience, domain-specific control beliefs, and attribution patterns. *Journal of Research in Science Teaching*, 35, 987-1013.
- Baumert, J., Stanat, P. & Demmrich, A. (2001). PISA 2000: Untersuchungsgegenstand, theoretische Grundlagen und Durchführung der Studie. In: Deutsches PISA-Konsortium (Hrsg.), *PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich* (15-68). Opladen: Leske & Budrich.
- Beaton, E. & Allen, N. (1992). Interpreting scales through scale anchoring. *Journal of Educational Statistics*, 17, 191–204.
- Bodemer, D., Ploetzner, R., Feuerlein, I., Spada, H. (2004). The active integration of information during learning with dynamic and interactive visualisations. *Learning and Instruction* 14, 325–341.
- Bond, T.G. & Fox, C.M. (2001). *Applying the Rasch Model: Fundamental measurement in the human sciences*. Mahwah, NJ: Erlbaum.
- Böttner, J., Maroska, R. & Olpp, A. (2005). *Schnittpunkt 2. Mathematik für Realschulen*. Stuttgart: Klett.

- Burnham, K.P. & Anderson, D.R. (2004). Multimodel inference: Understanding AIC and BIC in model selection. *Sociological Methods Research*, 33, 261-304.
- Carstensen, C. H., Ehmke, T. & Frey, A. (2008). *Ein mehrdimensionales Instrument zur Erhebung von Bildungsstandards in Mathematik*. Vortrag auf der 71. Tagung der AEPF, 27.08.08, Kiel.
- Csapó, B. (2004). Knowledge and competencies. In: J. Letschert (Ed.): *The integrated person. How curriculum development relates to new competencies* (35-49). Enschede: CIDREE/SLO.
- Desoete, A., Roeyers, H., & De Clercq, A. (2003). Can offline metacognition enhance mathematical problem solving? *Journal of Educational Psychology*, 95, 188-200.
- Draba, R. E. (1977). *The identification and interpretation of item bias*. Chicago: University of Chicago (Research Memorandum 26).
- Fischer, G. H. (1996). Unidimensional linear logistic rasch models. In: Linden, W. J. van der; Hambleton, R. K. (Eds.): *Handbook of modern item response theory*. New York, Berlin: Springer.
- Gagatsis, A. & Elia, I. (2004). The Effects of different modes of representation on mathematical problem solving. *Proceedings of the 28th Conference of the International Group for the Psychology of Mathematics Education*, 2, 447-454.
- Goldin, G.A. (1998): Representational systems, learning and problem solving in mathematics. *Journal of Mathematical Behavior*, 17 (2), 137-165.
- Harsch, C & Schröder, S. (2007). Testkonstruktion: C-Test. In: B. Beck. & E. Klieme. (Hrsg.): *Sprachliche Kompetenzen. Konzepte und Messung- DESI Studie*. (83-99). Weinheim: Beltz.
- Hartig, J. (2007). Skalierung und Definition von Kompetenzniveaus. In: B. Beck & E. Klieme (Eds.), *Sprachliche Kompetenzen. Konzepte und Messung - DESI-Studie* (72-82). Weinheim: Beltz.
- Hartig, J. & Höhler, J. (2008): Representation of competencies in multidimensional IRT models with within- and between-item multidimensionality. *Journal of Psychology* 216/2. 89-101.

- Hartig, J. & Frey, A. (2005). *Application of different explanatory item response models for model based proficiency scaling*. Paper presented at the 70th Annual Meeting of the Psychometric Society in Tilburg, July 5-8.
- Heller, A. & Perleth J. (2000). *Kognitiver Fähigkeitstest für 4. bis 12. Klassen, Revision (KFT 4-12+ R)*. Göttingen: Hogrefe.
- Herget, W., Malitte, E. & Richter, K. (2000): Funktionen haben viele Gesichter – auch im Unterricht. In: L. Flade & W. Herget (Hrsg.). *Mathematik lehren und lernen nach TIMSS: Anregungen für die Sekundarstufe* (115-124). Berlin: Volk und Wissen.
- Holling, H., Vock, M., & Wittmann, A. J. (2001). Wie erkennt man Hochbegabung bei jungen Erwachsenen? In: Deutsche Gesellschaft für das hochbegabte Kind e.V. (Hrsg.), *Im Labyrinth: Hochbegabte Kinder in Schule und Gesellschaft* (S. 16-21). Münster: LIT.
- Jäger, A.O., Süß, H.-M. & Beauducel, A. (1997). *Berliner Intelligenzstrukturtest Form 4 (BIS-T4)*. Göttingen: Hogrefe.
- Kaiser, F. G., & Wilson, M. (2004). Goal-directed conservation behavior: The specific composition of a general performance. *Personality and Individual Differences*, 36, 1531-1544.
- Klein, F., & Schimmack, R. (1907). *Vorträge über den mathematischen Unterricht an den höheren Schulen, Teil 1*. Von der Organisation des mathematischen Unterrichts. Leipzig.
- Kleine, M. (2005). Latent-Class-Analyse: Ein Bindeglied zwischen Empirie und Theorie zur quantitativen Erfassung mathematischer Leistungen. *Journal für Mathematik-Didaktik* 26 (2), 97–113.
- Klieme, E. & Hartig, J. (2008). Kompetenzkonzepte in den Sozialwissenschaften und im erziehungswissenschaftlichen Diskurs. In: M. Prenzel, I. Gogolin & H.-H. Krüger (Eds.), *Kompetenzdiagnostik: Sonderheft der Zeitschrift für Erziehungswissenschaft* (15-24). 8/2007. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Klieme, E. & Leutner, D. (2006). Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen. *Zeitschrift für Pädagogik*. 52 (6). 876-903.
- Klieme, E., Avenarius, H., Blum, W., Döbrich, P., Gruber, H., Prenzel, M., et al. (2003). *Zur Entwicklung nationaler Bildungsstandards*. Bonn: Bundesministerium für Bildung und Forschung.

- Klieme, E., Hartig, J., Kenk, M. & Fleischer, J. (2009). *Hinweise zu Konzepten der Kompetenzmodellierung*. Internes Paper im DFG-SSP „Kompetenzmodelle“.
- Klieme, E., Leutner, D. & Wirth, J. (Hrsg.) (2005). *Problemlösekompetenz von Schülerinnen und Schülern. Diagnostische Ansätze, theoretische Grundlagen und empirische Befunde der deutschen PISA-2000-Studie*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- KMK (Ständige Konferenz der Kultusminister der Länder der Bundesrepublik Deutschland) (Ed.). (2005a). *Bildungsstandards der Kultusministerkonferenz. Erläuterungen zur Konzeption und Entwicklung*. München: Wolters Kluwer.
- KMK (Ständige Konferenz der Kultusminister der Länder der Bundesrepublik Deutschland) (Ed.) (2005b). *Hessen, Lehrplan Mathematik. Gymnasialer Bildungsgang. Jahrgangsstufen 5G bis 12G*. Verfügbar unter <http://www.kultusministerium.hessen.de> [25.10.2009].
- Kozma, R. B. & Russell, J. (1997). Multimedia and understanding: Expert and novice responses to different representations of chemical phenomena. *Journal of Research in Science Teaching* 34, 593–619.
- Krüger, K. (1999). *Erziehung zum funktionalen Denken. Zur Begriffsgeschichte eines didaktischen Prinzips*. Frankfurt: Logos Verlag.
- Leuders, T. & Prediger, S. (2005). Funktioniert`s? – Denken in Funktionen. *Praxis der Mathematikdidaktik*. 2, (47), 1-7.
- Malle, G. (2000): *Zwei Aspekte von Funktionen: Zuordnung und Kovariation. Mathematik lehren*. Heft 103, Seelze: Friedrich-Verlag
- Matt, D. (2008). *Analyse und Konstruktion von Lern- und Leistungsaufgaben im Kompetenzbereich „proportionale und lineare Zusammenhänge“*. Wissenschaftliche Hausarbeit. Pädagogische Hochschule Freiburg.
- Mayer, R. E. & Hegarty, M. (1996). The process of understanding mathematics problems. In R. J. Sternberg & T. Ben-Zeev (Eds.), *The nature of mathematical thinking* (29 – 53). Hillsdale, NJ: Erlbaum.
- McClelland, D. C. (1973). Testing for competence rather than for „intelligence“. *American Psychologist*, 28, 1-14.

- Ministerium für Kultus, Jugend und Sport Baden-Württemberg (2004). *Bildungsplan für das Gymnasium der Normalform*. Verfügbar unter: <http://www.ls-bw.de/allg/lp> [25.10.2009].
- Moosbrugger, H. & Hartig, J. (2003). Faktorenanalyse. In K. Kubinger & R. Jäger (Hrsg.): *Schlüsselbegriffe der Psychologischen Diagnostik* (137-145). Weinheim: Psychologie Verlags Union.
- Muthén, B. & Muthén, L. (2000). Integrating person-centered and variable-centered analysis: Growth mixture modeling with latent trajectory classes. *Alcoholism: Clinical and Experimental Research*, 24, 882-891.
- Muthén, L. K. & Muthén, B. O. (2006). *Mplus Version 4.1. Statistical analysis with latent variables. User's guide*. Los Angeles, CA: Muthén & Muthén.
- Niss, M. (2003): *Mathematical Competencies and the Learning of Mathematics: The Danish KOM Project*. In: A. Gagatsis & S. Papastavridis (Hrsg.). 3rd Mediterranean Conference on Mathematical Education. Athens – Hellas 3-5 January 2003. (115-124). Athens: The Hellenic Mathematical Society.
- OECD (Organisation for Economic Co-operation and Development) (Ed.). (2005). *PISA 2003 Technical Report*. Paris: OECD Publishing.
- Pesonen M., Ehmke T. & Haapasalo L. (2005). Solving mathematical problems with dynamical sketches: a study on binary operations. In: *Problem Solving in Mathematics Education*. Proceedings of the Promath Meeting June 30-July 2, in Lahti (127–140).
- Prenzel, M., Baumert, J., Blum, W., Lehmann, R., Leutner, D., Neubrand, M., Pekrun, R., Rost, J. & Schiefele, U. (Hrsg.) (2005). *PISA 2003. Der zweite Vergleich der Länder in Deutschland – was wissen und können Jugendliche?* Münster: Waxmann.
- Rauch, D. & Hartig, J. (2007). Interpretation von Testwerten in der IRT. In H. Moosbrugger & A. Kelava (Hrsg.): *Test- und Fragebogenkonstruktion* (S. 240-250). Berlin: Springer.
- Reiss, K. (2009). *The impact of different dimensions of self-regulation in heuristic worked-out examples learning environment*. Presented at the EARLI Conference, Amsterdam, The Netherlands, 25.08.-29.08.2009.
- Rost, J. & Carstensen, C. H. (2002). Multidimensional Rasch measurement via item component models and faceted designs. *Applied Psychological Measurement*, 26, 42-56.
- Rost, J. (2003). Latent class analysis. In: R. Fernandez-Ballesteros (Ed.), *Encyclopedia of psychological assessment* (539-543). Thousand Oaks: Sage Publications.

- Rost, J., Carstensen, C. H., Bieber, G., Neubrand, M. & Prenzel, M. (2003). Naturwissenschaftliche Teilkompetenzen im Ländervergleich. In: Deutsches PISA-Konsortium (Hrsg.): *PISA 2000. Ein differenzierter Blick auf die Länder der Bundesrepublik Deutschland* (109-129). Opladen: Leske und Budrich.
- Rost, J. (2004). *Lehrbuch Testtheorie/ Testkonstruktion*. Bern: Verlag Hans Huber.
- Rupp, A. (2005). Quantifying subpopulation differences for a lack of invariance using complex examinee profiles: an exploratory multigroup approach using functional data analysis. *Educational Research and Evaluation*, 11, 71-97.
- Schnotz, W. (2005). An integrated model of text and picture comprehension. In: R.E. Mayer (Hrsg): *The Cambridge handbook of multimedia learning* (49-69). Cambridge: Cambridge University Press.
- Seidel, T. & Prenzel, M. (2008). Assessment in Large-Scale Studies. In: J. Hartig, E. Klieme & D. Leutner (Eds.), *Assessment of competencies in educational contexts: state of the art and future prospects* (279-304). Göttingen: Hogrefe & Huber.
- Seufert, T. (2003): Supporting coherence formation in learning from multiple representations. In: *Learning and Instruction* (13), 227–237.
- Snow, R. E. & Yalow, E. (1982). Education and intelligence. In: R.J. Sternberg (Hrsg.): *A handbook of human intelligence* (493–586). Cambridge: Cambridge University Press
- Swan, M. (1985). *The Language of Functions and Graphs*. Nottingham, UK: Shell Centre for Mathematical Education.
- Vermunt, J.K. (2004). Multilevel latent class models. *Sociological Methodology* 33, 213–239.
- Vermunt, J.K. (2008). Latent class and finite mixture models for multilevel data sets. *Methods in Medical Research* 17, 1, 33–51.
- Vollrath, H. J.(1989). Funktionales Denken. *Journal für Mathematikdidaktik*, 1, 3-37.
- von Davier, M. (2005). *A general diagnostic model applied to language testing data*. ETS Research Report 0x-2005.
- Walker, C. M. & Beretvas, S. N. (2003). Comparing multidimensional and unidimensional proficiency classifications: Multidimensional IRT as a diagnostic aid. *Journal of Educational Measurement*, 40, 255-275.

- Weinert, F. E. (2001). Concept of Competence: A Conceptual Clarification. In: D. S. Rychen & L. H. Salganik (Eds.), *Defining and Selecting Key Competencies* (45- 65). Seattle: Hogrefe & Huber.
- Wilson, M. (2003). On choosing a model for measuring. *Methods of Psychological Research Online*, 8, 1-22.
- Wilson, M. (2005). *Constructing measures. An item-response modelling approach*. Mahwah: Lawrence Erlbaum.
- Wilson, M., de Boeck, P. & Carstensen, C. (2008). Explanatory item response models: A brief introduction. In: J. Hartig, E. Klieme & D. Leutner (Eds.), *Assessment of competencies in educational contexts: state of the art and future prospects* (91-120). Göttingen: Hogrefe & Huber.
- Winter, H. (1995). Mathematikunterricht und Allgemeinbildung. In: *Mitteilungen der Gesellschaft für Didaktik der Mathematik*, 61. 37–46.
- Wirtz, M. & Caspar, F. (2002). *Beurteilerübereinstimmung und Beurteilerreliabilität: Methoden zur Bestimmung und Verbesserung der Zuverlässigkeit von Einschätzungen mittels Kategoriensystemen und Ratingskalen*. Göttingen: Hogrefe.
- Wu, M. & Adams, R. (2006). Modelling mathematics problem solving item responses using a multidimensional IRT model. *Mathematics Education Research Journal*, 18, 93-113.
- Wu, M. L., Adams, R. J. & Wilson, M. (2001). *ACER ConQuest version 2.0: generalised item response modelling software*. Camberwell, Victoria: ACER Press.

9 ANHANG

9.1	Itemkennwerte: Pilotierung	109
9.2	Itemkennwerte: Hauptstudie	111
9.3	Skalenkennwerte der Moderatorvariablen (Schülerfragebogen)	113

Itemkennwerte: Pilotierung

Item	δ_{uni}	SD	WMNSQ	T
1	1.600	0.506	1.08	0.4
2	0.950	0.434	1.08	0.5
3	-0.376	0.405	1.27	2.1
4	-0.899	0.435	1.02	0.2
5	-0.053	0.399	1.15	1.4
6	-0.716	0.422	1.04	0.3
7	-0.213	0.401	1.07	0.6
8	-0.543	0.412	0.97	-0.2
9	1.600	0.506	0.93	-0.1
10	0.265	0.401	0.87	-1.2
11	1.360	0.474	1.17	0.7
12	2.677	0.741	0.94	0.1
13	2.220	0.620	1.01	0.2
14	0.265	0.401	0.97	-0.2
15	0.594	0.412	1.13	0.9
16	1.146	0.451	0.97	-0.1
17	-1.095	0.452	1.02	0.2
18	0.106	0.399	0.97	-0.2
19	1.600	0.506	0.94	-0.1
20	0.950	0.434	0.83	-0.9
21	-0.213	0.401	1.00	0.0
22	-0.716	0.422	1.01	0.1
23	0.265	0.401	1.12	1.1
24	0.950	0.434	1.07	0.4
25	3.421	1.023	1.03	0.3
26	2.677	0.741	0.98	0.2
27	-0.716	0.422	0.94	-0.3
28	-0.213	0.401	1.01	0.1
29	1.600	0.506	0.91	-0.2
30	-0.053	0.399	1.06	0.6
31	1.600	0.506	0.99	0.1
32	1.360	0.474	1.00	0.1
33	-0.899	0.435	0.99	0.0
34	1.600	0.506	0.81	-0.6
35	0.428	0.405	0.89	-0.9
36	-0.439	0.438	1.37	2.2
37	-0.250	0.433	0.81	-1.3
38	-1.852	0.573	1.06	0.3
39	-0.063	0.431	0.76	-1.8
40	-1.289	0.496	0.99	0.0

Item	δ_{uni}	error	WMNSQ	T
41	0.919	0.470	0.83	-0.7
42	-1.289	0.496	0.92	-0.2
43	-0.063	0.431	0.75	-1.9
44	-0.838	0.457	0.67	-1.9
45	1.700	0.566	1.04	0.2
46	-1.289	0.496	0.88	-0.4
47	-0.838	0.457	0.86	-0.7
48	-2.218	0.642	0.87	-0.2
49	-0.838	0.457	0.78	-1.2
50	0.312	0.436	0.82	-1.1
51	1.150	0.492	1.06	0.3
52	1.700	0.566	1.23	0.7
53	-0.063	0.431	1.08	0.6
54	-0.250	0.433	0.83	-1.2
55	-1.852	0.573	0.94	-0.0
57	-0.250	0.433	0.91	-0.6
58	-0.250	0.433	0.82	-1.3
59	-1.852	0.573	1.09	0.4
60	-1.852	0.573	1.19	0.6
61	-1.852	0.573	0.86	-0.3
62	1.150	0.492	1.19	0.8
63	-1.289	0.496	0.78	-0.8
64	0.919	0.470	1.00	0.1
65	-0.439	0.438	0.74	-1.8
66	-1.289	0.496	0.77	-0.9
67	-1.054	0.474	0.72	-1.3
68	0.123	0.432	0.74	-1.9
69	0.505	0.444	1.27	1.5
70	0.123	0.432	0.97	-0.2
71	0.123	0.432	0.98	-0.1
72	1.405	0.522	1.03	0.2
73	0.123	0.432	0.99	0.0
74	2.057	0.634	1.31	0.7
75	1.150	0.492	1.32	1.2
76	0.706	0.454	0.82	-0.9
77	-0.063	0.431	0.88	-0.8
78	-0.063	0.431	0.89	-0.7

Itemkennwerte: Hauptstudie

Item	δ_{uni}	WMNSQ _{uni}	δ_{two}	WMNSQ _{two}	δ_{four}	WMNSQ _{four}
GG						
2 it_19bGGC	0,15	1,07	0,14	1,07	0,14	1,08
3 it_21bGGA	-1,6	0,92	-1,63	0,97	-1,64	0,93
6 it_33adGA	0,85	1,14	0,85	1,1	0,86	1,16
10 it_71cGNB	-0,98	1,07	-1,01	1,08	-1,01	1,05
15 it_18GGB	-0,89	1	-0,92	0,98	-0,93	0,97
16 it_23GGA	1,57	1,04	1,57	1,07	1,581	1,03
20 it_14aGGB	-1,18	0,97	-1,21	0,94	-1,22	0,95
23 it_17GGB	-0,2	0,92	-0,22	0,93	-0,23	0,9
24 it_20aGGA	-0,26	1,05	-0,27	1,03	-0,28	1,04
30 it_110bGGA	-0,3	1,04	-0,32	0,99	-0,33	1,03
35 it_11GGCa	0,87	0,95	0,88	0,99	0,87	0,96
36 it_11GGCb	0,56	0,98	0,56	0,95	0,55	0,95
37 it_21aGGA	0,35	0,97	0,35	0,97	0,34	0,94
51 it_20bGGA	-1	1,02	-1,01	1,06	-1,04	1,03
59 it_19aGGC	-0,2	0,99	-0,21	1,02	-0,21	1,02
60 it_15GGB	0,61	0,99	0,62	1,03	0,62	0,96
67 it_100GSA	-2,52	1,02	-2,54	1,04	-2,56	1,06
68 it_110aGGA	1,43	0,97	1,44	0,95	1,46	0,97
75 it_16aGGB	0,97	0,94	0,97	0,91	0,98	0,86
SG						
4 it_32aSGB	0,42	1,03	0,43	1,03	0,43	1,05
7 it_1aSGB	-1,68	0,94	-1,71	0,95	-1,73	0,96
8 it_106aSGB	1,09	0,99	1,1	0,94	1,11	0,99
9 it_35dSGB	4,57	0,99	4,61	0,98	4,63	0,93
11 it_2aGSB	0,08	1,04	0,08	1,03	0,08	1,05
12 it_5dSGB	-0,24	0,96	-0,25	0,98	-0,25	1,01
13 it_36aSGA	-0,74	1,1	-0,76	1,08	-0,77	1,08
14 it_36bSGA	-1,53	0,99	-1,56	1,01	-1,57	0,97
25 it_27acSGA	0,78	0,97	0,79	0,97	0,79	0,96
28 it_37cGSB	-1,54	1,05	-1,56	1,02	-1,56	1,03
29 it_37dGSB	-0,16	1,12	-0,16	1,14	-0,16	1,11
31 it_154SNG	1,77	0,98	1,79	1,01	1,8	1,01
32 it_72aSGB	1,62	0,92	1,63	0,97	1,65	0,97
38 it_24aSGA	-0,98	1,02	-1	0,96	-1	0,96
39 it_24bSGB	-0,17	0,94	-0,18	0,94	-0,17	0,89
40 it_28aGSB	-0,95	1,06	-0,96	1,03	-0,96	1
42 it_35bSGB	0,45	0,98	0,45	0,98	0,46	0,98
46 it_2bGSB	0,6	0,98	0,6	0,99	0,6	1
48 it_13aSGC	-0,81	1	-0,83	1,03	-0,84	1,03

53 it_107SGB	1,59	1,05	1,6	1,07	1,61	1,07
55 it_74SGBa	-0,21	0,97	-0,22	0,96	-0,22	0,94
56 it_74SGBb	-0,56	0,97	-0,58	0,98	-0,58	0,95
61 it_27cSGA	-0,57	1,04	-0,59	1,11	-0,61	1,11
71 it_12aSGB	-0,93	1	-0,95	0,99	-0,974	0,97
72 it_12bSGB	0,26	0,89	0,25	0,83	0,24	0,81
73 it_13bSGC	-0,67	1,01	-0,69	1,05	-0,71	1,02
74 it_30aSGB	-1,85	1,01	-1,88	0,97	-1,91	1
76 it_28bSGB	0,56	0,95	0,56	0,92	0,55	0,91
79 it_73SGB	-0,72	1,03	-0,73	1	-0,75	1,04
80 it_108GSB	-0,08	1,01	-0,09	0,96	-0,11	1,03
NN						
1 it_14bNNA	-1,06	0,9	-1,08	0,95	-1,2	0,92
21 it_109GSB	-1,38	1,02	-1,41	1,02	-1,53	1,03
22 it_7NNC	-0,32	1	-0,34	1	-0,36	0,95
44 it_103NNB	-0,63	1,04	-0,65	1,02	-0,69	1,04
49 it_16bNNB	0,32	1,07	0,32	1,02	0,36	1,04
62 it_31aNNC	-1,59	1,03	-1,6	1,07	-1,74	1,02
SN						
5 it_30bSNB	-1,4	0,95	-1,43	1	-1,5	1,05
18 it_35aSNB	-1,35	0,98	-1,37	1	-1,44	1
26 it_34aSNBa	-1,35	0,92	-1,38	0,92	-1,46	0,87
27 it_34aSNBb	-1,24	0,93	-1,26	0,92	-1,34	0,85
33 it_70SNC	0,18	1,05	0,18	0,99	0,17	1,1
34 it_1bSNB	-2,62	1,03	-2,66	1,03	-2,78	1,07
47 it_5cSNB	0,31	1,01	0,31	1,02	0,32	1,05
50 it_26bSNB	0,43	1,05	0,43	1,05	0,44	1,06
54 it_35cSNB	0,49	0,95	0,49	0,97	0,51	1
57 it_8abSNBa	-1,49	0,96	-1,53	0,93	-1,59	0,92
58 it_8abSNBb	-1,22	0,97	-1,26	0,94	-1,3	0,95
63 it_37aSNB	-1,9	1,06	-1,93	1,09	-2,01	1,12
64 it_37bSNB	-0,29	1,03	-0,31	1,06	-0,32	1,12
65 it_72bSNB	2,38	0,95	2,39	0,97	2,5	0,98
69 it_5abSNBa	-1,32	1,06	-1,34	1,04	-1,39	1,11
70 it_5abSNBb	0,03	0,98	0,03	0,98	0,05	0,99

Skalenkennwerte der Moderatorvariablen (Schülerfragebogen)

Selbstwirksamkeit Mathematik (PISA, 2000)

Variablen-name	Item	m	SD	r _{it}
SW1	Anhand eines Zugfahrplans ausrechnen, wie lange die Fahrt von einem Ort zu einem anderen dauern würde	2,55	1,21	0,77
SW2	Ausrechnen, wie viel billiger ein Fernseher bei 30 % Rabatt wäre.	2,41	1,10	0,78
SW3	Ausrechnen, wie viele Quadratmeter Fliesen du brauchst, um einen Fußboden mit bekannten Maßen damit auszulegen.	2,50	1,05	0,77
SW4	Diagramme in Zeitungen verstehen (z.B. zu Wahlergebnissen).	2,52	1,16	0,77
SW5	Eine Gleichung wie $3x + 5 = 17$ lösen.	2,57	1,27	0,78
SW6	Auf einer Karte im Maßstab 1:10.000 die tatsächliche Entfernung zwischen zwei Orten bestimmen.	2,62	0,966	0,88
Skala	Cronbachs $\alpha = 0,83$ m = 2,53 SD = 0,005 min = 1 max = 4 Theoretisches Max. = 4 N = 700			

Selbstreflektiertes Lernen (PISA, 2000)

Variablen- name	Item	m	SD	r _{it}
SLK1	Wenn ich lerne, zwingen mich zu prüfen, ob ich das Gelernte auch behalten habe.	2,69	1,03	0,48
SLK2	Wenn ich lerne, versuche ich beim Lesen herauszufinden, was ich noch nicht richtig verstanden habe.	2,59	0,85	0,48
SLK3	Wenn ich lerne, passe ich genau auf, dass ich das Wichtigste behalte.	3,12	0,86	0,42
SLK4	Wenn ich lerne und etwas nicht verstehe, suche ich nach zusätzlicher Information, um das Problem zu klären.	2,68	1,03	0,44
SLK5	Wenn ich lerne, überlege ich mir zuerst, was genau ich lernen muss.	2,69	1,03	0,56
Skala	Cronbachs $\alpha = 0,53$ m = 2,87 SD = 0,10 min = 1 max = 4 Theoretisches Max. = 4 N = 708			

Elaborationsstrategien (SLE 1-4):

Variablen-name	Item	m	SD	r _{it}
SLE1	Wenn ich lerne, versuche ich, neuen Stoff mit Dingen zu verbinden, die ich in anderen Fächern gelernt habe.	1,91	0,93	0,48
SLE2	Wenn ich lerne, überlege ich, inwiefern die Information im wirklichen Leben nützlich sein könnte.	2,03	1,02	0,66
SLE3	Wenn ich lerne, versuche ich den Stoff besser zu verstehen, indem ich Verbindungen zu Dingen herstelle, die ich schon kenne	2,26	0,89	0,46
SLE4	Wenn ich lerne, überlege ich, wie der Stoff mit dem zusammenhängt, was ich schon gelernt habe.	2,26	0,92	0,59
Skala	Cronbachs $\alpha = 0,62$ m = 2,12 SD = 0,03 min = 1 max = 4 Theoretisches Max. = 4 N = 709			

Anstrengungsbereitschaft (SLA 1-3):

Variablen-name	Item	m	SD	r _{it}
SLA1	Wenn ich lerne, arbeite ich auch dann weiter, wenn der Stoff schwierig ist.	2,86	0,90	0,69
SLA2	Wenn ich lerne, versuche ich mein Bestes zu geben, um mir das Wissen und die Fertigkeiten anzueignen.	3,01	0,96	0,52
SLA3	Wenn ich lerne, gebe ich mein Bestes.	3,05	0,84	0,45
Skala	Cronbachs $\alpha = 0,66$ m = 2,97 SD = ,009 min = 1 max = 4 Theoretisches Max. = 4 N = 703			

Wiederholungsstrategien (SLW 1-2):

Variablen-name	Item	m	SD
SLW1	Wenn ich lerne, versuche ich alles auswendig zu lernen, was drankommen könnte.	2,24	1,03
SLW2	Wenn ich lerne, präge ich mir alles Neue so ein, dass ich es aufsagen kann.	2,32	1,01
Skala	Cronbachs $\alpha = 0,51$ m = 2,28 SD = 0,003 min = 1 max = 4 Theoretisches Max. = 4 N = 727		

Interesse Mathematik (SLI 1-3)

Variablen-name	Item	m	SD	r _{it}
SLI1	Wenn ich mich mit Mathematik beschäftige, vergesse ich manchmal alles um mich herum.	2,51	0,97	0,81
SLI2	Mathematik ist mir persönlich wichtig.	2,54	0,94	0,64
SLI3	Die Beschäftigung mit Mathematik macht mir Spaß.	2,51	0,96	0,58
Skala	Cronbachs $\alpha = 0,77$ m = 2,52 SD = 0,00 min = 1 max = 4 Theoretisches Max. = 4 N = 706			

Angst Mathematik (PISA, 2000)

Variablen-name	Item	m	SD	r _{it}
A1	Vor Mathematik habe ich Angst.	2,51	1,26	0,73
A2	In Mathematik mache ich mir Sorgen, dass ich weniger als die anderen verstehe	2,49	1,07	0,75
A3	Beim Erledigen meiner Mathematikhausaufgaben mache ich mir Sorgen, ob ich das überhaupt allein schaffe.	2,48	1,06	0,69
Skala	Cronbachs $\alpha = 0,80$ m = 2,49 SD = 0,00 min = 1 max = 4 Theoretisches Max. = 4 N = 715			

Extrinsische Motivation (Bruder, 2005)

Variablen-name	Item	m	SD	r _{it}
EM1	Ich lerne für die Schule, weil ich sie erfolgreich abschließen möchte	2,43	1,22	0,90
EM2	Ich lerne für die Schule, weil ich gute Leistungen zeigen möchte	2,46	1,04	0,93
EM3	Ich lerne für die Schule, weil ich erfolgreich sein möchte	2,45	1,13	0,91
EM4	Ich lerne für die Schule, weil ich bei den Klassenarbeiten möglichst gut abschneiden möchte.	2,48	1,08	0,91
Skala	Cronbachs $\alpha = 0,93$ m = 2,45 SD = 0,00 min = 1 max = 4 Theoretisches Max. = 4 N = 712			

Schulunlust (Angstfragebogen für Schüler, AFS)

Variablen- name	Item	m	SD	r_{it}
SU1	Ich bin doch recht froh, dass ich noch zur Schule gehen kann. (-)	2,59	0,95	0,74
SU2	Schon der Gedanke an die Schule macht mich morgens oft missmutig.	2,58	1,52	0,78
SU3	Es wäre schön, wenn ich nicht mehr zur Schule gehen bräuchte.	2,47	0,99	0,73
SU4	Es gibt in der Schule eigentlich nur wenige Dinge, die einem wirklich Spaß machen.	2,54	0,96	0,73
SU5	Ich gehe gern zur Schule. (-)	2,59	0,89	0,69
SU6	Es ist doch schön, wieder in die Schule gehen zu können, auch wenn die Ferien noch so schön waren. (-)	2,55	1,04	0,73
Skala	Cronbachs $\alpha = 0,77$ m = 2,556 SD = 0,002 min = 1 max = 4 Theoretisches Max. = 4 N = 677			

Instuktionsverhalten Lehrkräfte (Bruder, 2008)

Variablen-name	Item	m	SD	r _{it}
IdL1	Unsere Lehrerin / unser Lehrer kann gut erklären.	2,49	1,055	0,78
IdL2	Unsere Unterrichtsstunden sind klar strukturiert	2,51	0,91	0,77
IdL3	Unsere Lehrerin / unser Lehrer gibt klare Anweisungen	2,49	0,98	0,69
Skala	Cronbachs $\alpha = 0,82$ m = 2,50 SD = 0,00 min = 1 max = 4 Theoretisches Max. = 4 N = 705			

Kompetenzunterstützung Lehrer (Bruder, 2008)

Variablen-name	Item	m	SD	r _{it}
Kus1	Die Lehrerin / der Lehrer traut mir zu, dass ich Mathe kann.	2,46	1,00	0,83
Kus2	Die Lehrerin/der Lehrer traut uns auch schwierige Aufgaben zu.	2,42	1,03	0,82
Kus3	Die Lehrerin/der Lehrer traut uns zu, selbstständig zu arbeiten.	2,39	1,11	0,82
Kus4	Die Lehrerin / der Lehrer kann mein Leistungsvermögen richtig einschätzen.	2,44	0,90	0,84
Kus5	Im Mathematikunterricht kann ich Hilfe bekommen, wenn ich sie benötige.	2,39	1,10	0,82
Skala	Cronbachs $\alpha = 0,86$ m = 2,42 SD = 0,001 min = 1 max = 4 Theoretisches Max. = 4 N = 677			

Kognitive Selbstständigkeit (Coaktiv)

Variablen- name	Item	m	SD	r _{it}
KS1	Unser Mathematiklehrer/ unsere Mathematiklehrerin lässt unterschiedliche Lösungswege von Aufgaben vergleichen und bewerten.	2,47	1,051	0,741
KS2	Im Mathematikunterricht lässt uns der Lehrer/ die Lehrerin auch einmal mit unseren eigenen Vermutungen in die Irre gehen, bis wir es selbst merken.	2,41	0,88	0,76
KS3	Im Mathematikunterricht akzeptiert der Lehrer/ die Lehrerin manchmal auch Fehler und lässt uns damit weitermachen, bis wir selbst sehen, dass etwas nicht stimmt.	2,42	0,90	0,76
KS4	Unser Mathematiklehrer/ unsere Mathematiklehrerin stellt auch Aufgaben, bei denen man mehrere Lösungswege zeigen muss.	2,49	0,89	0,78
KS5	Unser Mathematiklehrer/ unsere Mathematiklehrerin lässt uns häufiger unsere Gedankengänge genau erklären.	2,44	0,95	0,77
KS6	Unser Mathematiklehrer/ unsere Mathematiklehrerin verlangt häufiger, dass wir unsere Arbeitsschritte ausführlich begründen.	2,52	1,11	0,76
KS7	Bei unserem Mathematiklehrer/ unserer Mathematiklehrerin kann ich Aufgaben so lösen, wie ich es persönlich für richtig halte.	2,35	0,90	0,78
KS8	Bei unserem Mathematiklehrer/ unserer Mathematiklehrerin kann ich zum Lösen schwieriger Aufgaben meine eigenen Strategien einsetzen.	2,34	0,98	0,76
Skala	Cronbachs $\alpha = 0,79$ m = 2,43 SD = ,004 min = 1 max = 4 Theoretisches Max. = 4 N = 665			

DANKSAGUNG

Mein besonderer Dank gilt meinen Betreuern Prof. Dr. Markus Wirtz und Prof. Dr. Timo Leuders für die Betreuung der Arbeit, die Vergabe des Themas und dafür, dass sie mir die Möglichkeit boten, am DFG-Schwerpunktprogramm „Kompetenzmodelle“ teilzunehmen.

Weiterhin danke ich allen Schülerinnen und Schülern, die sorgfältig die Tests bearbeitet haben. Auch ihren Lehrer/innen sei dafür gedankt, dass sie Unterrichtsstunden zur Verfügung stellten und Fragebögen bearbeiteten. Ohne sie wäre diese Arbeit nicht möglich gewesen.

Für die Hilfe bei der Datenerhebung in den Schulen und bei der Dateneingabe möchte ich mich bei allen Tutor/innen bedanken, insbesondere bei Jonas Hummel sowie Katja Kreft, Carola Bernack, Dominik Matt und Anja Schmälzle.

Der Deutschen Forschungsgemeinschaft (DFG) danke ich für die Finanzierung der Arbeit, für die zahlreichen Fortbildungsveranstaltungen sowie für Vernetzung mit anderen Forschergruppen der empirischen Bildungsforschung im Rahmen des Schwerpunktprogramms „Kompetenzmodelle“. Den Sprechern des Schwerpunktprogramms Prof. Dr. Eckeard Klieme und Prof. Dr. Detlev Leutner sei herzlich für die Einrichtung des Schwerpunktprogramms gedankt. Weiterer Dank gilt den anderen Doktorand/innen des SPP, namentlich Samuel Greiff, Nora Neuenhaus und Birgit Harks für den fachlichen und persönlichen Austausch.

Ein besonderes Dankeschön geht an meinen Schreibtischnachbarn Marco Petrucci und meine Flurnachbarn Evelyn Müller und Enrico Prinz für die positive Arbeitsatmosphäre und die anregenden Diskussionen.

Ganz herzlich bedanken möchte ich mich schließlich bei meinem Mann Carsten Habeck, meinen Eltern Horst und Ulrike Bayrhuber, meiner Schwester Monika Bayrhuber sowie meinen Freundinnen Barbara Schürmann und Beate Schlimm für ihre Unterstützung, aus der ich viel Kraft für die Fertigstellung dieser Arbeit geschöpft habe.