pedocs
Open Access Erziehungswissenschaften

DIPF

Praetorius, Anna-Katharina; Greb, Karina; Lipowsky, Frank; Gollwitzer, Mario

Lehrkräfte als Diagnostiker. Welche Rolle spielt die Schülerleistung bei der Einschätzung von mathematischen Selbstkonzepten?

Journal for educational research online 2 (2010) 1, S. 121-144

urn:nbn:de:0111-opus-45707

in Kooperation mit / in cooperation with:

WAXMANN VERLAG GMBH
Münster · New York · München · Berlin

http://www.waxmann.com

Nutzungsbedingungen / conditions of use

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

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

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.
By using this particular document, you accept the above-stated conditions of use.

Kontakt / Contact:

peDOCS
Deutsches Institut für Internationale Pädagogische Forschung (DIPF)
Mitglied der Leibniz-Gemeinschaft
Informationszentrum (IZ) Bildung
Schloßstr. 29, D-60486 Frankfurt am Main
E-Mail: pedocs@dipf.de
Internet: www.pedocs.de

Journal for Educational Research Online
Journal für Bildungsforschung Online
Volume 2 (2010), No. 1, 121–144
© 2010 Waxmann

Anna-Katharina Praetorius, Karina Greb, Frank Lipowsky
& Mario Gollwitzer

Lehrkräfte als Diagnostiker – Welche Rolle spielt die Schülerleistung bei der Einschätzung von mathematischen Selbstkonzepten?

Zusammenfassung

Die vorliegende Studie untersucht, ob und in welchem Ausmaß Lehrkräfte zur Einschätzung der mathematischen Fähigkeitsselbstkonzepte ihrer Schüler/-innen deren Mathematikleistungen einbeziehen. Der Fragestellung wird anhand eines Strukturgleichungsmodells sowie eines Mehrebenenmodells nachgegangen, basierend auf einer Stichprobe von 36 Lehrkräften und 663 Erstklässler/-innen. Es zeigt sich, dass Lehrereinschätzungen des mathematischen Schülerselbstkonzepts stärker durch die Leistungen der Schüler als durch deren Fähigkeitsselbstkonzepte vorhergesagt werden. Ferner lassen sich bedeutsame Unterschiede zwischen Lehrern in Bezug auf die Vorhersage der Lehrereinschätzungen durch die Schülerselbsteinschätzungen aufweisen. Die Ergebnisse werden im Hinblick auf ihre theorie- und praxisbezogenen Implikationen diskutiert.

Schlagworte

Diagnostische Kompetenz, Urteilsgenauigkeit, mathematisches Fähigkeitsselbstkonzept, Mathematikleistung

Teachers' Diagnostic Perception – What Role Does Students' Grade Performance Play in the Teacher's Assessment of Mathematical Self-concepts?

Abstract

This study focuses on whether and to which extent teachers incorporate grade performance in mathematics in appraising the students' self-concepts of their own mathematical ability. Structural equation and multilevel modelling was used to investigate these questions based on 36 teachers and 663 first graders. Mathematical achievement predicted teacher ratings of students' mathematical self-concept to a greater degree than students' mathematical self-concepts did. However, the degree to which teacher ratings were predicted by pupils' ratings differed substantially across teachers. A discussion on the theoretical and practical implications of the results concludes this study.

Keywords
diagnostic competence, judgment accuracy, academic self-concept in mathematics, achievement in mathematics

Aktuellen Forschungsergebnissen zufolge verfügen Grundschullehrkräfte im Bereich des Fähigkeitsselbstkonzepts über eine eher unzureichend ausgeprägte diagnostische Kompetenz (vgl. z. B. Praetorius, Greb, Dickhäuser & Lipowsky, in Druck; Spinath, 2005). Ein Grund dafür könnte sein, dass Lehrkräfte eher die – im Mittel wohl salienteren – Leistungen von Schülerinnen und Schülern[1] als deren Fähigkeitsselbstkonzepte für ihre Einschätzung heranziehen. Ob sich diese Vermutung empirisch bestätigen lässt, soll in der vorliegenden Untersuchung überprüft werden.

1. Bedeutung des Untersuchungsgegenstandes

Diagnostische Kompetenz – definiert als die „Fähigkeit von Lehrkräften, Schülerinnen und Schüler sowie lern- und leistungsrelevante Sachverhalte zutreffend zu beurteilen“ (Schrader, 2009, S. 237) – zählt zu den zentralen Kompetenzen von Lehrkräften (vgl. Artelt & Gräsel, 2009; Standards für die Lehrerbildung der Kultusministerkonferenz, 2004; Weinert, Schrader & Helmke, 1990). Bezüglich der wissenschaftlichen Erforschung der diagnostischen Kompetenz von Lehrkräften steht bisher die Diagnose von Leistungen im Vordergrund (z. B. Schrader, 1989; Schrader & Helmke, 1987; Schrader, Helmke, Hosenfeld, Halt & Hochweber, 2006). Mittlerweile wird zudem vermehrt auf die Bedeutung adäquater Lehrereinschätzungen von nicht-kognitiven Schülermerkmalen – etwa motivationaler Merkmale – hingewiesen (vgl. Artelt & Gräsel, 2009; Spinath, 2005; Südkamp & Möller, 2009).

Ein bedeutendes motivationales Merkmal ist dabei das Fähigkeitsselbstkonzept. Dieses kann nach Schöne, Dickhäuser, Spinath und Stiensmeier-Pelster (2002) definiert werden als die „Gesamtheit der kognitiven Repräsentationen eigener Fähigkeiten in akademischen Leistungssituationen“ (S. 7). Dem Fähigkeitsselbstkonzept kommt aus mehreren Gründen eine zentrale Rolle für den Unterricht zu: Es beeinflusst das Erleben und Verhalten in Leistungssituationen in vielfältiger Weise und wirkt sich über affektiv-motivationale Variablen auch auf Leistungen aus (für weiterführende Informationen vgl. Möller & Trautwein, 2009; vgl. auch Dickhäuser, 2006). Zudem kommt einem positiven Selbstkonzept unabhängig von den Zusammenhängen mit anderen Faktoren aus pädagogischer Sicht ein nicht zu unterschätzender Eigenwert zu (vgl. Einsiedler, Martschinke & Kammermeyer, 2007; Trautwein, 2003). Dementsprechend findet sich in Lehrplänen unterschiedlicher Schularten und Bundesländer die Aufforderung, dass Lehrkräfte ihre Schüler

1 Aus Gründen der Lesbarkeit wird im Folgenden lediglich die männliche Form verwendet. Es sind jedoch stets beide Geschlechter gemeint.

in einer positiven Persönlichkeitsentwicklung unterstützen sollen (z. B. Bayerisches Staatsministerium für Unterricht und Kultus, 2000).

Es wird davon ausgegangen, dass diagnostisch kompetente Lehrkräfte ihr pädagogisches Handeln optimieren und damit die Persönlichkeits- und Lernentwicklung ihrer Schüler positiv unterstützen können (z. B. Helmke, 2009; Weinert & Schrader, 1986). Konkret lässt sich annehmen, dass Lehrkräfte, wenn sie die Fähigkeitsselbstkonzepte ihrer Schüler kennen, ihr Instruktions- und Rückmeldeverhalten eher an die Voraussetzungen der Schüler anpassen können. Insbesondere wenn Schüler über stark unrealistische, insbesondere unrealistisch niedrige Fähigkeitsselbstkonzepte verfügen, erscheinen entsprechende Interventionen – etwa in Form von Reattributionstrainings – sinnvoll (vgl. z. B. Dresel, 2001). Bei den genannten Punkten handelt es sich jedoch lediglich um Plausibilitätsannahmen: Zu tatsächlichen empirischen Wirkungen diagnostischer Kompetenz gibt es bislang kaum Untersuchungen (vgl. Schrader, 2009). Die wenigen vorhandenen deutschsprachigen Untersuchungen hierzu beziehen sich auf den Bereich der Schüler*leistungen*. So weist eine Studie von Hosenfeld (2009) darauf hin, dass es einen direkten Effekt der diagnostischen Kompetenz auf die Lernentwicklung der Schüler gibt, während eine Untersuchung von Helmke und Schrader (1987) aufzeigt, dass sich eine hohe diagnostische Kompetenz lediglich indirekt – nämlich in Verbindung mit einem ausreichenden Maß an unterrichtlichen Strukturierungs- und Unterstützungsmaßnahmen – positiv auf den Lernerfolg von Schülern auswirkt.

2. Akkuratheitskomponenten diagnostischer Kompetenz

Hohe positive Produkt-Moment-Korrelationen zwischen dem Lehrerurteil und einem geeigneten Kriterium gelten nach Schrader und Helmke (1987) als Indikator „für die diagnostische Kompetenz im eigentlichen Sinne“ (S. 35). Eine hohe positive Korrelation bedeutet, dass relative Unterschiede zwischen den beurteilten Personen im Lehrerurteil auch relativen Unterschieden im Kriterium (z. B. den Selbsteinschätzungen der Schüler) entsprechen. Aus diesem Grund wird die Korrelation zwischen Lehrerurteil und Kriterium auch als „Rangkomponente“ bezeichnet. Die Korrelation bezieht jedoch nicht mit ein, inwiefern Lehrer die „wahren“ Merkmalsausprägungen sowie die Unterschiedlichkeit zwischen denselben systematisch unter- oder überschätzen. Daher sollten neben der Korrelation zusätzlich die durchschnittliche Abweichung zwischen Lehrerurteilen und Kriterium („Niveaukomponente“) sowie der Unterschied zwischen der Streuung der Lehrerurteile und der Streuung der Kriteriumsvariablen („Differenzierungskomponente“) als weitere Akkuratheitsmaße mit angegeben werden.[2] Es werden im

2 Für weiterführende Informationen zu den Komponenten diagnostischer Kompetenz siehe z. B. Cronbach (1955), Greb (2009) oder Schrader und Helmke (1987).

Ergebnisteil folglich alle drei Komponenten aufgeführt. Die beiden letztgenannten Komponenten spielen für die vorliegende Fragestellung jedoch keine zentrale Rolle und werden daher aus Platzgründen in der Darstellung des Forschungsstands nicht berücksichtigt.

3. Gegenstands- und Domänenspezif tät von diagnostischer Kompetenz

Bisherige Forschungen zur diagnostischen Kompetenz von Lehrkräften zeigen, dass diese *gegenstandsspezifisch* (also abhängig davon, welches Schülermerkmal eingeschätzt wird) ausgeprägt ist. So können Lehrkräfte die Leistungen ihrer Schüler vergleichsweise gut einschätzen: Die durchschnittlichen Produkt-Moment-Korrelationen zwischen Schülerleistung und Lehrerurteil liegen im – nach Cohen (1992) – mittleren bis hohen Bereich. In der Metaanalyse von Hoge und Colardarci (1989) findet sich z. B. ein Median von $r = .66$ (Range: $.28 \leq r \leq .92$), in einer Untersuchung von Schrader und Helmke (1987) eine durchschnittliche Übereinstimmung von $r = .64$ und in einer Studie von Schrader und Kollegen (2006) eine mittlere Korrelation von $r = .51$. Auch die Einschätzung der allgemeinen Intelligenz von Schülern gelingt Lehrkräften moderat bis gut: Bei Spinath (2005) findet sich beispielsweise im Mittel eine Übereinstimmung zwischen Schülerintelligenz und Lehrereinschätzung von $r = .40$, bei Wild (1991) eine mittlere Korrelation von $r = .67$. Die Fähigkeitsselbstkonzepte von Schülern schätzen Lehrkräfte im Durchschnitt mäßig gut ein. Dies zeigen insbesondere Studien der Forschergruppe um Marsh (z. B. Marsh, 1990: $r = .30$; Marsh, Smith & Barnes, 1984: $r = .31$), aber auch deutschsprachige Untersuchungen wie z. B. diejenige von Spinath (2005), in der sich eine mittlere Korrelation von $r = .39$ findet. Die schulische Leistungsängstlichkeit von Schülern einzuschätzen scheint Lehrkräften nach bisherigen Forschungsbefunden hingegen äußerst schwer zu fallen. So findet sich z. B. bei Federer, Stüber, Margraf, Schneider und Herrle (2001) lediglich eine durchschnittliche Übereinstimmung von $r = .08$, Ergebnisse von Helmke und Fend (1981) weisen sogar nicht signifikant von Null verschiedene Korrelationen auf. *Systematische* Vergleiche dazu, wie zutreffend die gleichen Lehrpersonen unterschiedliche Merkmale von Schülern (z. B. Leistungen und Fähigkeitsselbstkonzepte) einschätzen, existieren unseres Wissens bislang für Grundschullehrkräfte nicht.

In Bezug auf die Frage, inwiefern die diagnostische Kompetenz von Lehrkräften systematisch über *unterschiedliche Schulfächer bzw. Domänen* hinweg variiert, zeigen sich in einer Studie von Lorenz und Artelt (2009) für die leistungsbezogene diagnostische Kompetenz sehr ähnliche Zusammenhangsmuster zwischen Lehrereinschätzungen und Schülerleistungen über unterschiedliche Fächer hinweg (z. B. für den ersten Messzeitpunkt ihrer Längsschnittuntersuchung: Wortschatz $r = .55$, Textverstehen $r = .60$, Arithmetik $r = .56$). Das bedeutet: Lehrpersonen gelingt es in allen drei untersuchten Domänen vergleichsweise gut, die leistungs-

bezogene Rangreihe ihrer Schüler einzuschätzen. In einer weiteren Untersuchung zur diagnostischen Kompetenz, hier bezogen auf das Fähigkeitsselbstkonzept von Schülern, konnte ebenfalls gezeigt werden, dass die Bereiche Mathematik und Lesen mit einer Produkt-Moment-Korrelation von $r = .55$ bzw. $r = .52$ zwischen Lehrereinschätzung und Schülermerkmal ähnlich gut eingeschätzt werden (Praetorius et al., in Druck). Anders verhält es sich für den Bereich Schreiben. Für diesen zeigt sich in der genannten Untersuchung mit $r = .25$ eine deutlich niedrigere Korrelation. In einem weiteren Schritt wurde zudem untersucht, ob Lehrkräfte, welche die Fähigkeitsselbstkonzepte ihrer Schüler in einem Bereich (z. B. Mathematik) gut einschätzen können, das auch in einem anderen Bereich (z. B. Lesen) können. Es zeigt sich, dass dies lediglich für die beiden Bereiche des Schriftspracherwerbs, Lesen und Schreiben, gilt: Je zutreffender Lehrkräfte die Fähigkeitsselbstkonzepte ihrer Schüler im Schreiben einschätzen, desto zutreffender fallen auch die Urteile über die Fähigkeitsselbstkonzepte der Schüler im Lesen aus und umgekehrt ($r = .38$, $p < .01$). Zwischen den Bereichen Schreiben und Mathematik sowie Lesen und Mathematik lassen sich hingegen keine solchen Zusammenhänge finden ($r = .04$ und $r = .14$, beide n. s.). Die genannten Befunde verweisen darauf, neben der leistungsbezogenen diagnostischen Kompetenz auch die diagnostische Kompetenz bezogen auf Fähigkeitsselbstkonzepte von Schülern stets nach Fächern bzw. Domänen differenziert zu untersuchen.

Für die vorliegende Untersuchung beziehen wir uns auf das Schulfach Mathematik.

4. Quellen von Lehrereinschätzungen

4.1 Ursachen für unerklärte Varianz: systematische und unsystematische Fehler

Die in der bisherigen Forschung gefundenen Korrelationen zwischen Lehrereinschätzungen und Fähigkeitsselbstkonzepten befinden sich – wie oben erläutert – im moderaten Bereich zwischen $.30 \leq r \leq .60$ (zsf. Spinath, 2005). Betrachtet man den Varianzanteil, der durch das von den Schülern selbstberichtete Fähigkeitsselbstkonzept an der Gesamtvarianz der Lehrereinschätzungen desselben aufgeklärt werden kann, stellt sich die Frage, worauf der verbleibende Anteil unerklärter Varianz zurückzuführen ist.

Schrader und Helmke (1990) unterscheiden im Bereich der leistungsbezogenen diagnostischen Kompetenz zwei Möglichkeiten zur Erklärung einer solchen niedrigen Varianzaufklärung: unsystematische sowie systematische Einflüsse. Unsystematische Einflüsse beziehen sich auf die Unreliabilität der Messung von Urteil und Kriterium.

Systematische Einflüsse hingegen sind den Autoren zufolge solche, die auf die Einschätzungen selbst zurückzuführen sind: In diesem Fall wird die Akkuratheit

der Lehrereinschätzungen dadurch reduziert, dass die Lehrereinschätzungen etwas anderes messen als das, was eigentlich gemessen werden soll (in unserem Fall also das Fähigkeitsselbstkonzept der Schüler).

Untersuchungen, die sich damit beschäftigen, in welchem Ausmaß in ein Urteil neben dem eigentlich einzuschätzenden Merkmal weitere Merkmale – wie z. B. unterschiedliche Personenmerkmale des zu Beurteilenden – einfließen, finden sich im Bereich der Lehrerforschung kaum (Cooksey, Freebody & Davidson, 1986; vgl. auch Cooksey, 1996).

4.2 Einbezug weiterer Schülermerkmale im Leistungsbereich

Eine der wenigen diesbezüglichen Untersuchungen stammt von Schrader und Helmke (1990). Die Autoren stellten fest, dass in die Einschätzung von Schülerleistungen durch Lehrkräfte neben den Schülerleistungen auch die Intelligenz sowie das Fähigkeitsselbstkonzept der Schüler einfließen. Eine mögliche Erklärung für den Einbezug der Intelligenz ist laut Schrader und Helmke (1990), dass Intelligenz und Leistung auch empirisch einen Zusammenhang aufweisen und daher von Lehrern schwer zu unterscheiden sind. Ein Grund, warum Lehrkräfte Fähigkeitsselbstkonzepte von Schülern in ihr Urteil über deren Leistungen einbeziehen, könnte den Autoren zufolge in der je nach Höhe des Fähigkeitsselbstkonzepts unterschiedlichen Selbstdarstellung von Schülern begründet liegen. Ein weiterer Erklärungsansatz für den Einbezug von Intelligenz und Fähigkeitsselbstkonzept, der durch vertiefende Analysen von Schrader und Helmke (1990) gestützt wird, geht davon aus, dass die beiden Merkmale Aspekte von fachlicher Leistung abdecken, die mit dem verwendeten Leistungstest nur partiell bzw. nicht reliabel erfasst wurden.

4.3 Einbezug weiterer Schülermerkmale im Bereich des Fähigkeitsselbstkonzepts

Welche Informationen Lehrkräfte zur Einschätzung von Fähigkeitsselbstkonzepten heranziehen, wurde bislang kaum untersucht. Eine Ausnahme bildet eine Untersuchung von Pohlmann, Möller und Streblow (2004). Unter Rückgriff auf das Internal/External-frame-of-reference-Modell (z. B. Marsh, 1984), welches postuliert, dass *Schüler* zur Einschätzung ihrer eigenen Fähigkeiten sowohl soziale als auch dimensionale Vergleiche vornehmen,[3] vermuteten die Autoren, dass auch *Lehrkräfte* zur Einschätzung der Fähigkeitsselbstkonzepte von Schülern auf soziale und dimensionale Vergleichsinformationen rekurrieren. Pohlmann und

3 Soziale Vergleiche (= external frame) meinen dabei die interpersonalen Vergleiche der Leistungen verschiedener Schüler, dimensionale Vergleiche (= internal frame) die intrapersonalen Vergleiche der Leistung eines Schülers in einem Fach mit der Leistung desselben Schülers in einem anderen Fach.

Kollegen stellten jedoch fest, dass Lehrkräfte keine dimensionalen Vergleiche zur Einschätzung von Fähigkeitsselbstkonzepten vornehmen, wohl aber soziale Vergleiche:[4] In ihrer Untersuchung findet sich ein stark positiver Einfluss der Mathematiknote auf die mathematikbezogene Fähigkeitsselbstkonzepteinschätzung durch die Mathematiklehrkräfte, während der Effekt der Deutschnote auf die Einschätzung des mathematischen Fähigkeitsselbstkonzepts durch die Mathematiklehrkräfte nicht signifikant ist. Die Autoren erklären dieses Ergebnis folgendermaßen: „Zur Einschätzung anderer Personen (...) werden eher leicht zugängliche Informationen genutzt, solange sie für die Beurteilung hinreichend sind. Dieses Vorgehen erscheint durchaus sinnvoll, da der soziale Vergleich allein bereits eine recht zufrieden stellende Schätzung zulässt“ (ebd., S.9).

Anknüpfend an die Interpretation der oben genannten Studie durch Pohlmann und Kollegen, wonach Lehrkräfte insbesondere jene Informationen nutzen, die ihnen leicht zugänglich sind, lässt sich annehmen, dass Lehrkräfte bei der Einschätzung von Fähigkeitsselbstkonzepten verstärkt auf die Leistungen der Schüler zurückgreifen – auf Daten also, mit deren Erfassung sie zum einen vergleichsweise häufig konfrontiert sind und die zum anderen, zumindest in weiterführenden Schulen, einen annähernd akzeptablen Indikator für das Fähigkeitsselbstkonzept bilden (z. B. Schilling, Sparfeldt & Rost, 2006).

Denkbar ist zudem, dass Lehrkräfte in ihre Einschätzung von Fähigkeitsselbstkonzepten auch deshalb in einem hohen Ausmaß die Leistungen der Schüler einbeziehen, weil sie in ihrer Ausbildung eher wenig über Fähigkeitsselbstkonzepte von Kindern an sich, die Entwicklung derselben und die Zusammenhänge von Fähigkeitsselbstkonzepten und Lernverhalten erfahren.

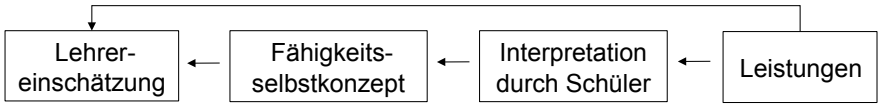
Geht man davon aus, dass Lehrkräfte vor allem soziale *Leistungs*vergleiche zur Einschätzung von Fähigkeitsselbstkonzepten verwenden, sollten sie in der Folge Fähigkeitsselbstkonzepte, die lediglich in einem geringen Zusammenhang mit den korrespondierenden Leistungen stehen, deutlich schlechter einschätzen können. Solche Fähigkeitsselbstkonzepte finden sich verstärkt zu Beginn der Grundschule: Die meisten Schüler schätzen hier ihre eigenen Fähigkeiten sehr hoch ein (z. B. Fend & Stöckli, 1996; Helmke, 1998; Kammermeyer & Martschinke, 2004). Dementsprechend finden sich in der ersten Klasse deutlich geringere Korrelationen zwischen dem Fähigkeitsselbstkonzept und den korrespondierenden Leistungsmaßen als in weiterführenden Schulen ($.18 \leq r \leq .27$ bei Bouffard, Marcoux, Vezeau & Bordeleau, 2003 sowie Kammermeyer & Martschinke, 2006; nicht signifikant bei Chapman & Tunmer, 1997). Legt man die Überlegungen von Pohlmann und Kollegen (2004) zugrunde, sollten Grundschullehrkräfte, die in ersten Klassen unterrichten, daher über eine deutlich geringere diagnostische Kompetenz verfügen als Lehrkräfte weiterführender Schulen. Spinath (2005) fasst

4 Das gleiche Befundmuster zeigt sich bei Untersuchungen von Marsh und Kollegen (z. B. Marsh, 1986; Marsh, Smith & Barnes, 1984). Ein gegenläufiges Ergebnis findet sich in einer experimentellen Untersuchung von Dickhäuser (2005): Die an der Studie teilnehmenden Studenten verwendeten auch dimensionale Vergleiche, wenn ihnen die entsprechenden Leistungsinformationen zur Verfügung gestellt wurden.

den diesbezüglichen Forschungsstand zusammen und kommt zu dem Fazit, dass die Übereinstimmungen zwischen den Fähigkeitsselbstkonzepten der Lernenden und deren Einschätzungen durch die Lehrpersonen tatsächlich mit steigender Klassenstufe höher sind. Trotzdem liegen sie durchgehend im moderaten Bereich – soziale Vergleiche der Leistung können demnach nicht die alleinige Grundlage der Lehrereinschätzungen sein.

Auf Grundlage der aufgeführten Überlegungen gehen wir von folgenden Annahmen zum Zusammenhang zwischen dem Schülerfähigkeitsselbstkonzept, der Schülerleistung und der vom Lehrer vorgenommenen Einschätzung des Schülerfähigkeitsselbstkonzepts aus (vgl. Abbildung 1):

Abbildung 1: Annahmen zum Zusammenhang zwischen den Leistungen der Schüler, der Interpretation von Leistungssituationen, dem Fähigkeitsselbstkonzept der Schüler sowie der Lehrereinschätzung des Fähigkeitsselbstkonzepts



Auf Seiten der Schüler entstehen Fähigkeitsselbstkonzepte – wie aus der Forschung hinlänglich bekannt (z. B. Dickhäuser, 2006; Spinath, 2004) – aufgrund vielfältiger Interpretationen von Leistungssituationen. Diese Interpretationen beruhen u.a. auf indirekten Rückmeldungen (basierend auf den besagten sozialen oder dimensionalen Vergleichen) und direkten Rückmeldungen (von Lehrkräften, Eltern und Mitschülern) (vgl. Moschner & Dickhäuser, 2006). In der vorliegenden Untersuchung wird davon ausgegangen, dass Lehrkräfte, wenn sie die Fähigkeitsselbstkonzepte von Schülern einschätzen wollen, keine Differenzierung wie die obige vornehmen, sondern die Ausprägungen der Fähigkeitsselbstkonzepte eher holistisch und unter Anwendung von heuristischen Strategien erschließen (vgl. das Konzept der bounded rationality; z. B. Fuchs-Heinritz, 2007; Gigerenzer & Selten, 2001).

Leistung ist eine zentrale Dimension von Schulerfolg (z. B. Wiater, 2006) und wird von den Lehrkräften somit aufgrund ihrer Bedeutung vermutlich stark wahrgenommen. Wir postulieren daher in Anlehnung an Pohlmann und Kollegen (2004), dass neben dem Fähigkeitsselbstkonzept die Schülerleistung einen zusätzlichen direkten Einfluss auf die Einschätzung von Fähigkeitsselbstkonzepten hat (*Hypothese 1*).

Da das Fähigkeitsselbstkonzept im Unterricht im Gegensatz zur Leistung nur indirekt ersichtlich ist und Leistung in der Unterrichtswahrnehmung der meisten Lehrer einen zentraleren Stellenwert einnimmt – und damit als salienterer Reiz fungiert als das Fähigkeitsselbstkonzept – gehen wir davon aus, dass die Leistung zudem einen stärkeren Einfluss auf die Fähigkeitsselbstkonzepteinschätzung hat als das Fähigkeitsselbstkonzept der Lernenden selbst (*Hypothese 2*).

4.4 Unterschiede zwischen Lehrkräften im Einbezug von Leistung und Fähigkeitsselbstkonzept

Bisherige Studien deuten auf große interindividuelle Unterschiede in der diagnostischen Kompetenz von Lehrkräften hin (Rangkomponente für den Leistungsbereich z. B. Schrader & Helmke, 1987, $.02 \leq r \leq .88$ und für den Bereich des Fähigkeitsselbstkonzepts z. B. Spinath, 2005, $-.39 \leq r \leq .82$). Diese Unterschiede sind laut Schrader (2009) auf urteilserleichternde bzw. -erschwerende Kontextmerkmale (z. B. Leistungsniveau und -heterogenität der Klasse, vgl. dazu auch Karing, 2009) sowie auf Kompetenzunterschiede zwischen Lehrkräften zurückzuführen.

Dabei hängt die Urteilsgenauigkeit laut Schrader vor allem davon ab, in welchem Umfang das Urteil der jeweiligen Lehrkraft auf validen Indikatoren basiert. Dies setzt implizit voraus, dass Lehrkräfte sich erstens dahingehend unterscheiden, welche Indikatoren sie in ihr Urteil einbeziehen und zweitens, wie sie diese Indikatoren gewichten. Empirische Studien hierzu liegen unseres Wissens bislang nicht vor.

Da Grundschullehrkräfte nahezu täglich mit der Diagnostik und Beurteilung von Schülerleistungen konfrontiert sind und diese somit in irgendeiner Weise bei Lehrkräften mental repräsentiert sein dürften, lässt sich annehmen, dass Lehrkräfte sich kaum darin unterscheiden, wie stark sie die Leistungen der Schüler in ihre Einschätzung von Fähigkeitsselbstkonzepten einbeziehen. Hingegen ist zu vermuten, dass Lehrer sich im Wissen über Fähigkeitsselbstkonzepte und in ihrer mentalen Repräsentation derselben deutlich unterscheiden. Dies wiederum sollte dazu führen, dass Lehrkräfte darin variieren, in welchem Ausmaß sie die Fähigkeitsselbstkonzepte der Schüler einbeziehen. In der vorliegenden Studie soll untersucht werden, ob sich Lehrkräfte wie vermutet darin unterscheiden, in welchem Ausmaß sie die Fähigkeitsselbstkonzepte in ihre Einschätzungen einbeziehen (*Hypothese 3.1*), nicht jedoch darin, in welchem Ausmaß sie die Leistungen der Schüler einbeziehen (*Hypothese 3.2*).

5. Zusammenfassung der Hypothesen

H1: Lehrkräfte ziehen zur Einschätzung von mathematischen Fähigkeitsselbstkonzepten sowohl die Mathematikleistung als auch das mathematische Fähigkeitsselbstkonzept heran.

H2: Der Einfluss der Mathematikleistung auf die Einschätzung des mathematischen Fähigkeitsselbstkonzepts ist größer als derjenige des mathematischen Fähigkeitsselbstkonzepts selbst.

H3.1: Es finden sich keine signifikanten Unterschiede zwischen Lehrkräften dahingehend, in welchem Ausmaß die Mathematikleistungen der Schüler in die Lehrereinschätzungen des Fähigkeitsselbstkonzepts einfließen.

H3.2: Lehrkräfte unterscheiden sich signifikant in dem Ausmaß, in dem sie die mathematischen Fähigkeitsselbstkonzepte ihrer Schüler in ihre Einschätzungen einbeziehen.

6. Methode

6.1 Stichprobe

Die Stichprobe der Arbeit basiert auf im Rahmen der Studie PERLE (*Persönlichkeits- und Lernentwicklung von Grundschulkindern*) erhobenen Daten (vgl. Lipowsky, Faust & Greb, 2009). An der Studie nahmen insgesamt 38 Klassen aus 20 Schulen teil. Die Schulen befinden sich in Sachsen, Thüringen, Berlin und Mecklenburg-Vorpommern. In die vorliegende Untersuchung wurden die Daten derjenigen Lehrpersonen einbezogen, von denen auch Einschätzungen der Schülerfähigkeitsselbstkonzepte vorliegen. Dabei handelt es sich um 36 Klassenlehrkräfte (35 Lehrerinnen, 1 Lehrer) sowie deren 663 Erstklässler. Die Berufserfahrung der Lehrkräfte variierte zwischen 1 und 35 Jahren (SD = 10.9) und lag im Mittel bei 16 Jahren. Das Alter der Schüler lag zwischen 5.7 und 9.7 Jahren (M = 7.5 Jahre). Die Klassenstärke betrug im Mittel 19 Schüler (SD = 3.58; min. = 13; max. = 26).

6.2 Instrumente und Durchführung

6.2.1 Mathematisches Fähigkeitsselbstkonzept der Schüler

Das mathematische Fähigkeitsselbstkonzept wurde am Ende des ersten Schuljahres (Juni 2007) gruppenweise anhand eines Fragebogens erhoben, der u.a. 6 Items zum mathematischen Selbstkonzept umfasst.[5] Die Items haben ein dreistufiges Antwortformat (Beispielitem: „Fällt dir das Rechnen leicht oder schwer?“ mit den Antwortstufen „leicht“ (3), „weder leicht noch schwer“ (2), „schwer“ (1)). Die interne Konsistenz (Cronbachs α) war zufriedenstellend (α = .83). Die durchschnittliche Höhe des Fähigkeitsselbstkonzepts lag bei M = 2.39 (SD = .43).

6.2.2 Mathematische Kompetenz der Schüler

Die allgemeine mathematische Kompetenz der Schüler wurde ebenfalls im Rahmen einer 40-minütigen Gruppenerhebung am Ende des ersten Schuljahres erfasst. Die Schüler bearbeiteten einen Test mit 44 Items, dessen Schwerpunkt auf ein-

5 Der Fragebogen umfasst insgesamt 24 Items, die sich neben zwei Beispielitems auf die vier Selbstkonzeptfacetten Mathematik (6 Items), Schreiben (4 Items), Lesen (4 Items) sowie kreative Tätigkeiten (8 Items) beziehen (vgl. Poloczek, 2007; Poloczek, Greb & Lipowsky, eingereicht).

fachen arithmetischen Operationen liegt (Eigenentwicklung, vgl. Greb, Lipowsky & Faust, i. V.). Die dichotom kodierten Antworten wurden im Anschluss Raschskaliert. Vier dieser Skalen können wiederum als Subdimensionen der arithmetischen Kompetenz bezeichnet werden (vgl. Padberg, 2007): Fähigkeit zur Lösung von Additionsaufgaben (19 Items), Vorläuferfähigkeiten zur Multiplikation (9 Items), Fähigkeit zur Lösung von Subtraktionsaufgaben (15 Items) und Rechnen mit Geldbeträgen (5 Items). Die EAP/PV-Reliabilität der Subdimensionen liegt zwischen .75 und .84 (Reliabilität allgemeine Additionsfähigkeit = .87; Reliabilität Vorläuferfähigkeiten zur Multiplikation = .84; Reliabilität allgemeine Subtraktionsfähigkeit = .87; Reliabilität Rechnen mit Geld = .75).

6.2.3 Lehrereinschätzung des Fähigkeitsselbstkonzepts

Die Lehrkräfte bearbeiteten am selben Tag der Schülererhebung einen Fragebogen, der u.a. die Einschätzung des Fähigkeitsselbstkonzepts der von ihnen unterrichteten Schüler in den verschiedenen Bereichen anhand eines Globalurteils umfasste. Damit die Lehrkräfte sich ein Bild von dem einzuschätzenden Gegenstand machen konnten – mit dem Ziel, die Validität der Einschätzungen zu erhöhen – wurden den Lehrkräften die Schülerselbstkonzeptitems vorgelegt. Zudem wurde darauf hingewiesen, dass die Fähigkeitsselbstkonzepte durch Mittelwertbildungen der Schüleritems berechnet werden.

Den Lehrkräften standen zur Einschätzung ihrer Schüler sechs Antwortmöglichkeiten zur Verfügung („sehr gering“ (1), „gering“ (2), „eher gering“ (3), „eher hoch“ (4), „hoch“ (5) und „sehr hoch“ (6)). Der Mittelwert der Lehrereinschätzung lag bei $M = 4.03$ ($SD = 1.20$).

7. Ergebnisse

7.1 Deskriptive Statistiken zu den drei Akkuratheitskomponenten

Die gemittelte Rangkomponente über alle Lehrkräfte hinweg weist einen Wert von $r = .55$ ($p < .01$) und eine Standardabweichung von $SD = .21$ auf (Range: $.00 \leq r \leq .89$). Der Mittelwert der Niveaukomponente liegt bei $M = -0.16$ bei einer Standardabweichung von $SD = 0.17$ (Range: $-0.50 \leq M \leq 0.20$). Für die Differenzierungskomponente findet sich ein Mittelwert von $M = 1.15$ mit einer Standardabweichung von $SD = 0.32$ (Range: $0.53 \leq M \leq 1.98$).[6]

6 Die Rangkomponente (Idealausprägung: 1) einer Lehrperson wird ermittelt, indem pro Klasse die Fähigkeitsselbstkonzepte der Schüler und die Einschätzungen derselben durch die Lehrperson korreliert werden. Die mittlere Rangkomponente über die gesamte Stichprobe setzt sich aus diesen klassenweise berechneten Produkt-Moment-Korrelationen zusammen. Diese werden vor ihrer Mittelung Fisher-Z-transformiert. Die Niveaukomponente (Idealausprägung: 0) wird durch die klassenweise Differenz zwischen dem

7.2 Einf uss von Leistung und Fähigkeitsselbstkonzept auf die Lehrereinschätzung

Die Überprüfung von Hypothese 1 und 2 erfolgt anhand eines Strukturgleichungsmodells (mit Mplus 5.1; Muthén & Muthén, 1998-2007). In die Analysen wurden dabei lediglich diejenigen Schülerwerte einbezogen, für welche auch Lehrerurteile vorliegen. Der vorliegenden Mehrebenenstruktur der Daten (Schachtelung von Schülern in Klassen bzw. Lehrkräften) wurde durch statistische Korrektur der Standardfehler unter Berücksichtigung der Intraklassenkorrelation Rechnung getragen.

Das Strukturmodell enthält die Lehrereinschätzung des mathematischen Fähigkeitsselbstkonzepts als endogene Variable sowie die Leistung der Schüler in Mathematik und das mathematische Fähigkeitsselbstkonzept der Schüler als exogene Variablen. Das Messmodell wurde auf Seiten der exogenen Variablen mit sechs Items zur Messung des mathematischen Fähigkeitsselbstkonzepts sowie mit vier Subskalen zur Messung der allgemeinen mathematischen Fähigkeit modelliert. Auf Seiten der endogenen Variablen wurde die schülerweise Einschätzung des Fähigkeitsselbstkonzepts anhand eines Globalurteils mit einem Item erfasst.

Zur Berechnung des Modells wurde der MLR-Schätzer (maximum likelihood restricted) verwendet. Dieser ist robuster gegenüber Verletzungen der Normalverteilungsvoraussetzung und kommt in der Folge zu einer verlässlicheren Schätzung der Standardfehler (Muthén & Muthén, 1998–2007).

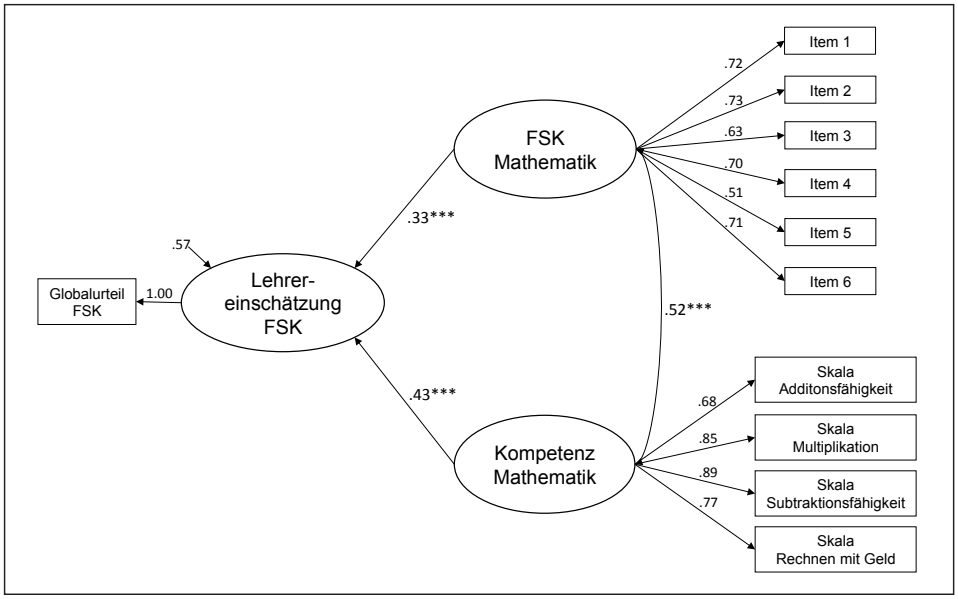
Das aufgestellte Modell verfügt über eine gute Modellanpassung ($\chi^2_{(.001;\ 42,\ N=663)}$ = 118.30; CFI = .97; RMSEA = .05; SRMR = .03). Insgesamt lassen sich durch das Modell 43 % der Varianz in den Lehrereinschätzungen aufklären.

Es zeigt sich, dass sowohl die Mathematikleistung der Schüler (β = .43, p < .001) als auch das Fähigkeitsselbstkonzept der Schüler (β = .33, p < .001) einen Einfluss auf die Lehrereinschätzung haben (s. auch Abbildung 2). Die Leistung der Schüler stellt dabei den erklärungsstärkeren Faktor dar: Im Vergleich zu einem Modell mit gleichgesetzten Pfadkoeffizienten verfügt ein Modell mit variablen Pfadkoeffizienten über eine signifikant bessere Anpassung an die Daten (χ^2-Differenzentest mit Satorra-Bentler-Skalierung: χ^2 = 19.61; p < .001). Hypothesen 1 und 2 können somit bestätigt werden.

---

mittleren Lehrerurteil und dem Mittelwert der Schülerwerte bestimmt. Die Differenzierungskomponente (Idealausprägung: 1) berechnet sich als Quotient der klassenweisen Streuung der Lehrereinschätzung und der tatsächlichen Streuung. Für weiterführende Analysen bezogen auf die Komponenten, z. B. zu den Zusammenhängen derselben, siehe Praetorius et al. (in Druck).

Abbildung 2: Strukturgleichungsmodell zum Zusammenhang zwischen Schülerleistung, Fähigkeitsselbstkonzept und der Lehrereinschätzung des Fähigkeitsselbstkonzepts



Anmerkungen. Bei den angegebenen Werten handelt es sich um standardisierte Werte; FSK = Fähigkeitsselbstkonzept; *** $p < .001$.

7.3 Unterschiede zwischen Lehrkräften im Einbezug von Leistung und Fähigkeitsselbstkonzept

Anhand des Statistikprogramms *HLM 6* (Raudenbush, Bryk, Cheong & Congdon, 2004) wurde überprüft, inwiefern Lehrkräfte sich darin unterscheiden, wie stark sie in ihre Einschätzung des Fähigkeitsselbstkonzepts die Leistung sowie das Fähigkeitsselbstkonzept einbeziehen (vgl. Hypothesen 3.1 und 3.2). Die Modellierung als Zweiebenenmodell wurde gewählt, da es sich bei dem vorliegenden Datensatz um eine hierarchische Datenstruktur handelt, bei der Schüler in Klassen geschachtelt sind (vgl. z. B. Snijders & Bosker, 1999).

Um die Frage nach Unterschieden zwischen Lehrkräften beantworten zu können, wurde ein Random-Coefficient-Modell spezifiziert, bei dem die Residualwerte auf Ebene 2 (u_{0j}, u_{1j} und u_{2j}) als Zufallsvariablen modelliert sind. Im Modell werden zur Erklärung der Lehrerurteile (y_{ij}: Einschätzung von Schüler i durch Lehrkraft j) lediglich Prädiktoren auf Schülerebene (in unserem Fall das Fähigkeitsselbstkonzept x_{1ij} sowie die Schülerleistung x_{2ij}), nicht jedoch Prädiktoren auf Klassen- bzw. Lehrkraftebene einbezogen. Das Modell lässt sich durch folgende Gleichungen darstellen:

Modell auf Ebene 1: $y_{ij} = \beta_{0j} + \beta_{1j} \cdot x_{1ij} + \beta_{2j} \cdot x_{2ij} + e_{ij}$

Modell auf Ebene 2:

$$\beta_{0j} = \gamma_{00} + u_{0j}$$
$$\beta_{1j} = \gamma_{10} + u_{1j}$$
$$\beta_{2j} = \gamma_{20} + u_{2j}$$

Die Schätzung der Regressionsgewichte β_{1j} und β_{2j} sowie des Achsenabschnittsparameters β_{0j} erfolgt im Rahmen von Mehrebenenmodellen klassen- bzw. lehrerspezifisch. Die Residuen e_{ij} repräsentieren dabei ausschließlich Abweichungen innerhalb der jeweiligen Klasse, die Residuen u_{0j}, u_{1j} und u_{2j} hingegen Abweichungen zwischen Klassen bzw. zwischen Lehrern. Die Residualvariablen auf Ebene 2 dürfen miteinander kovariieren; der Kovarianzmatrix wird keine Struktur auferlegt.

Zur Überprüfung der Frage, inwiefern Lehrkräfte sich darin unterscheiden, wie stark ihre Urteile durch das Fähigkeitsselbstkonzept der Schüler und durch die Leistungen der Schüler vorhergesagt werden, wurde überprüft, in welchem Ausmaß die random slopes (u_{1j} und u_{2j}) variieren. Die Lehrereinschätzungen, Fähigkeitsselbstkonzepte und Schülerleistungen gehen dabei als manifeste Variablen (Mittelwerte der oben genannten Indikatoren für die jeweilige Variable) in die HLM-Analyse ein. Sie wurden zudem vor den Berechnungen z-standardisiert.

Bei der Schätzung des Modells anhand des Restricted-Maximum-Likelihood-Verfahrens (RML) ergibt sich für den random slope des Fähigkeitsselbstkonzepts eine Varianz von $\sigma^2_{u1j} = .02$ ($\chi_{35}^2 = 53.25$; $p = .03$) und für denjenigen der Schülerleistung ebenfalls eine Varianz von $\sigma^2_{u2j} = .02$ ($\chi_{35}^2 = 45.28$; $p = .11$), die jedoch nicht signifikant ist. Während sich Lehrkräfte also dahingehend unterscheiden, in welchem Ausmaß sie bezogen auf die gesamte Klasse das Fähigkeitsselbstkonzept in ihre Einschätzung desselben einbeziehen, zeigen sich für den Einbezug der Leistung keine signifikanten Unterschiede zwischen den Lehrkräften.

Es muss jedoch einschränkend darauf hingewiesen werden, dass der Chi-Quadrat-Wert auch für die Varianz des Regressionsgewichts der Leistung (σ^2_{u2j}) nicht niedrig ist. Es kann also nicht ausgeschlossen werden, dass es lediglich auf die Stichprobe zurückzuführen ist, dass der Wert nicht signifikant wird. Der Befund sollte demnach in einer zukünftigen Studie auf seine Robustheit hin überprüft werden.

8. Diskussion

Aktuell wird vermehrt darauf hingewiesen, dass sich Untersuchungen zur diagnostischen Kompetenz nicht lediglich auf die Untersuchung der Genauigkeit von Lehrereinschätzungen beschränken, sondern darüber hinaus Erkenntnisse z. B. über Bedingungen, Stabilität und Auswirkungen von diagnostischer Kompetenz anstreben sollten (z. B. Helmke, Hosenfeld & Schrader, 2004; Schrader, 2009; Schrader, in Druck). Dieser Forderung wird im vorliegenden Artikel insofern

Rechnung getragen, als untersucht wird, in welchem Ausmaß sich die Lehrereinschätzung der mathematischen Fähigkeitsselbstkonzepte von Schülern nicht nur durch die Fähigkeitsselbstkonzepte selbst sondern auch durch die Leistungen der Schüler speist.

8.1 Leistung als Grundlage der Einschätzung von Fähigkeitsselbstkonzepten

Im Rahmen der vorliegenden Arbeit konnte gezeigt werden, dass Lehrkräfte in ihre Einschätzung von Fähigkeitsselbstkonzepten sowohl die Schülerleistungen als auch die Fähigkeitsselbstkonzepte einbeziehen, die Leistungen dabei in einem größeren Ausmaß.

Das Ergebnis, dass Lehrkräfte zur Beurteilung neben dem einzuschätzenden Merkmal ein weiteres Merkmal heranziehen, steht in Einklang mit Befunden aus der Lehr-Lernforschung, die zeigen, dass Lehrkräfte nicht distinkt zwischen verschiedenen Konstrukten unterscheiden: Wie schon erwähnt konnten z. B. Schrader und Helmke (1990) zeigen, dass Lehrkräfte zur Einschätzung von Schülerleistungen deren Intelligenz, insbesondere aber auch deren Fähigkeitsselbstkonzepte einbeziehen.

Erklären kann man den hier vorliegenden Befund auf dreierlei Art und Weise: Erstens weisen Leistung und Fähigkeitsselbstkonzept einen empirischen Zusammenhang auf und sind daher schwer unterscheidbar. Wenn Lehrkräfte Leistungen zur Beurteilung von Fähigkeitsselbstkonzepten heranziehen, sind sie daher nicht zwingend schlechte Diagnostiker. Zum zweiten sind Fähigkeitsselbstkonzepte deutlich schwieriger zu beobachten als schulische Leistungen. Es ist also nicht verwunderlich, dass Lehrkräfte auf die Leistungen der Schüler zurückgreifen, wenn sie deren Fähigkeitsselbstkonzepte einschätzen sollen. Eine dritte Ursache könnte in mangelndem Wissen von Lehrkräften über das Selbstkonzept und insbesondere dessen Entwicklung begründet sein.

8.2 Die Rolle des Fähigkeitsselbstkonzepts für die Lehrereinschätzungen

Über die Leistung hinaus beziehen Lehrkräfte aber auch die Fähigkeitsselbstkonzepte der Schüler in ihre Einschätzungen ein. Das deutet darauf hin, dass Lehrkräfte ihre Einschätzungen – wie eingangs schon vermutet – nicht nur auf Grundlage sozialer Leistungsvergleiche treffen. Dies ist positiv zu sehen: Diagnostische Sensitivität für lern- und leistungsrelevante Schülermerkmale gilt als wichtig für einen adaptiven Unterricht (vgl. Schrader, in Druck) und ist, gerade weil motivationale Merkmale wie das Fähigkeitsselbstkonzept nicht direkt ersichtlich sind, nicht selbstverständlich.

Relativiert wird dieser positiv zu wertende Befund durch die vorhandene Varianz zwischen den Lehrern: In der vorliegenden Studie konnte unter Berück-

sichtigung der oben formulierten methodischen Einschränkung gezeigt werden, dass sich Lehrkräfte nicht statistisch signifikant dahingehend unterscheiden, in welchem Ausmaß sie die Leistung der Schüler in ihre Einschätzung einbeziehen, wohl aber darin, in welchem Ausmaß sie das Fähigkeitsselbstkonzept einbeziehen. Das bedeutet: Zur Einschätzung der Fähigkeitsselbstkonzepte rekurrieren die von uns untersuchten Lehrkräfte in einem sehr ähnlichen Ausmaß auf Schülerleistungen, aber in unterschiedlichem Ausmaß auf die tatsächlichen Fähigkeitsselbstkonzepte der Lernenden.

8.3 Erklärungsmöglichkeiten von Unterschieden zwischen Lehrern

In zukünftigen Studien sollten Variablen einbezogen werden, welche dazu beitragen, die Unterschiede zwischen Lehrkräften in der Stärke des Einbezugs des Fähigkeitsselbstkonzepts zu erklären. Dabei sind auf Lehrkraft- bzw. Klassenebene drei potentielle Gruppen von Varianzquellen möglich, auf die Unterschiede zurückgeführt werden können:

So könnte die vorhandene Varianz auf Ebene 2 zunächst einmal auf *tatsächliche Klassenunterschiede* in Bezug auf das Fähigkeitsselbstkonzept zurückzuführen sein. Dies würde bedeuten, dass Lehrkräfte sich deswegen darin unterscheiden, in welchem Ausmaß sie Fähigkeitsselbstkonzepte in ihre Einschätzung einbeziehen, weil sich die Klassen bezogen auf die durchschnittliche Ausprägung des Fähigkeitsselbstkonzepts unterscheiden.

Eine zweite mögliche Varianzquelle bilden *Klassenvariablen* wie z. B. die Klassengröße, die Heterogenität der Merkmalsausprägung oder der Anteil der Schüler mit deutlich von den korrespondierenden Leistungen abweichenden Fähigkeitsselbstkonzepten. Aus Plausibiliätsgründen kann in etwa bezüglich der Klassengröße angenommen werden, dass Lehrkräfte in kleineren Klassen – aufgrund des höheren durchschnittlichen Zeitanteils pro Schüler – individueller auf diese eingehen und dieselben daher auch besser einschätzen können. Die empirische Forschungslage zu Effekten der Klassengröße bezogen auf diverse Kriterien ist jedoch als äußerst widersprüchlich zu bezeichnen (vgl. Blatchford, 2009; Ingenkamp, Petillon & Weiß, 1985). So zeigt z. B. eine schweizerische Studie mit Grund- und Sekundarschulklassen von Brühlwiler und Blatchford (in Druck), dass Lehrkräfte in kleineren Klassen nicht adaptiver unterrichten als in größeren.

Die dritte mögliche Gruppe an Einflussfaktoren sind *Lehrkraftvariablen* wie in etwa die Berufserfahrung, das Lehrerwissen über Selbstkonzepte oder aber generelle Urteilstendenzen von Lehrkräften (z. B. Tendenz zur Mitte). Für die Berufserfahrung in etwa könnte – ebenfalls aus Plausibilitätsgründen – ein Zusammenhang erwartet werden. Empirisch zeigt sich jedoch oftmals kein Effekt der Berufserfahrung auf die Höhe der diagnostischen Kompetenz (z. B. für den Leistungsbereich: Schrader, 1989; für den Bereich kognitiver Schülerfähigkeiten: Wild & Rost, 1995).

Betrachtet man zudem die Residualvarianz auf Ebene 1 (e_{ij}), lässt sich feststellen, dass ein großer Teil der unerklärten Varianz auf dieser Ebene zu verorten ist. Auch unter Berücksichtigung der Messfehlervarianz, die ebenfalls hier eingeht, kommt vermutlich schülerbezogenen Variablen auf Ebene 1 als Einflussfaktoren in Form von Kovariaten oder Moderatoren auf die Einschätzung von Fähigkeitsselbstkonzepten durch die Lehrkräfte eine hohe Bedeutung zu. Dies in weiteren Studien näher zu untersuchen ist ein interessanter Ansatz, der im Bereich der Forschung zu diagnostischer Kompetenz bislang so noch nicht verfolgt wird.

9. Implikationen für Forschung und Praxis

9.1 Durchführung prozessbezogener Untersuchungen

Was im vorliegenden Artikel ausgeklammert bleibt, ist die Frage, ob Lehrkräfte, die in einem stärkeren Ausmaß das Fähigkeitsselbstkonzept einbeziehen, tatsächlich auch über eine höhere diagnostische Kompetenz im Sinne der drei Akkuratheitskomponenten verfügen. Dies ist zu vermuten, allerdings bislang empirisch noch nicht überprüft. Sollte sich zeigen, dass dies der Fall ist, wären Mikroanalysen des Urteilsprozesses ebendieser Lehrkräfte von hohem Interesse. So könnte man Informationen über die Art und Weise der internen Verarbeitung von Informationen erhalten. Solche Untersuchungen wurden bislang kaum durchgeführt. Dies erscheint jedoch – unter der Voraussetzung, dass ein stärkerer Einbezug des Fähigkeitsselbstkonzepts tatsächlich auch mit einer höheren diagnostischen Kompetenz einhergeht – äußerst gewinnbringend: Die Expertise derjenigen Lehrkräfte, die vor allem das Fähigkeitsselbstkonzept einbeziehen, könnte so dazu genutzt werden, ein besseres Verständnis davon zu entwickeln, wie Einschätzungen von Fähigkeitsselbstkonzepten vonstattengehen können.

9.2 Identif kation von Diagnosemöglichkeiten

Sichtet man die bisherige Literatur, finden sich erstaunlich wenige Hinweise darauf, wie Lehrkräfte die Fähigkeitsselbstkonzepte ihrer Schüler erschließen können. In der Regel wird lediglich darauf hingewiesen, dass Fähigkeitsselbstkonzepte aus dem Verhalten der Schüler erschlossen werden müssen – anhand welcher Indikatoren dies geschehen kann, wird dabei nicht thematisiert. Gerade weil Fähigkeitsselbstkonzepte nicht direkt beobachtbar sind, sondern aus dem Verhalten indirekt erschlossen oder aber explizit erfragt werden müssen, sind Erhebungsverfahren wichtig, die Lehrkräfte kriteriengeleitet darin unterstützen, Fähigkeitsselbstkonzepte zu erfassen. Bislang existieren vor allem formelle Verfahren zur Diagnose von Fähigkeitsselbstkonzepten, die für den unterrichtlichen Gebrauch jedoch in der Regel zu aufwändig sind. Für die Entwicklung von in den Unterricht integrierbaren informellen Diagnoseverfahren könnten Erkenntnisse

aus den oben erwähnten Befragungen derjenigen Lehrkräfte genutzt werden, die in einem starken Maß die Fähigkeitsselbstkonzepte ihrer Schüler in die Einschätzung derselben einbeziehen.

9.3 Unterrichtsbezogene Interventionen

Wie schon angesprochen, schätzen viele Schüler ihre eigenen Fähigkeiten zu Beginn der Grundschule sehr hoch ein. Dies ist entwicklungspsychologisch gesehen normal (z. B. Harter, 2006). Es gibt jedoch auch Schüler, die sich auch hier schon stark unterschätzen. Um einer negativen Persönlichkeits- und Lernentwicklung entgegenzuwirken, kann daher der Einsatz entsprechender Interventionen sinnvoll sein.[7] Interventionen setzen dabei eine akkurate Diagnose von Fähigkeitsselbstkonzepten voraus. Lehrkräfte dürften sich jedoch nur dann bemühen, Fähigkeitsselbstkonzepte adäquat zu diagnostizieren und ihre eigene Diagnosekompetenz weiterzuentwickeln, wenn sie wissen, wie eine Umsetzung von Interventionen im eigenen Unterricht möglich ist. Bislang existieren vor allem Interventionsmaßnahmen, die für in diesem Bereich unerfahrene Lehrkräfte ohne Expertenhilfe kaum in den Unterricht integrierbar sind (z. B. computerbasierte Reattributionstrainings; Dresel & Ziegler, 2006). Zudem sind Reattributionstrainings laut Ziegler und Schober (2001) erst für Schüler ab einem Alter von etwa zehn Jahren geeignet. Bei jüngeren Schülern kann nicht davon ausgegangen werden, dass diese distinkt zwischen Fähigkeit und Anstrengung differenzieren sowie über Wissen darüber verfügen, dass Fähigkeiten modifizierbar sind. Es besteht somit deutlicher Entwicklungsbedarf bezüglich Interventionsmaßnahmen, die von Lehrkräften leicht im Unterricht einsetzbar sind und insbesondere solchen Maßnahmen, die sich für Schüler im Grundschulbereich eignen.

Das Ausmaß, in dem Lehrkräfte solche Interventionsmaßnahmen kennen und schätzen, sollte zudem positiv mit der Akkuratheit der Einschätzung von Selbstkonzepten korreliert sein. Diese Vermutung lässt sich aus Studien ableiten, die sich auf die soziale Kognitionsforschung beziehen (z. B. Krolak-Schwerdt, Böhmer & Gräsel, 2009; Krolak-Schwerdt & Rummer, 2005). In diesen Studien zeigt sich, dass erfahrene Lehrkräfte personenbezogene Informationen je nach Verarbeitungsziel differenziell verarbeiten (Stichwort: Lehrer als „flexible Denker“). Dabei wird zwischen dem Ziel der Eindrucksbildung und jenem der Vorhersage unterschieden. Ersteres steht in Verbindung mit einer kategorienbasierten Verarbeitung, bei der kaum spezifische Personencharakteristika einbezogen werden, letzteres mit einer merkmalsbasierten Verarbeitung, bei der diese Charakteristika explizit einbezogen werden. Verfolgen Lehrkräfte das Ziel,

7 Solchen Interventionen kommt auch auf dem Hintergrund, dass vor allem Mädchen ihre schulischen Kompetenzen in Mathematik und den Naturwissenschaften deutlich ungünstiger einschätzen als es deren Leistungen erwarten lassen (Bos, Bonsen, Baumert, Prenzel, Selter, Walther, 2007; Whitehead, 1996; Gabriel, Mösko & Lipowsky, in Druck), eine erhebliche Bedeutung zu.

Schüler mit niedrigen Selbstkonzepten durch intervenierende Maßnahmen in der Entwicklung eines positiveren Selbstkonzepts zu unterstützen, sollten sie daher eher darum bemüht sein, die Selbstkonzepte ihrer Schüler genau vorherzusagen. Im regulären Unterricht hingegen werden selbstkonzeptbezogene Informationen vermutlich – wenn überhaupt – eher mit dem Ziel der Eindrucksbildung wahrgenommen und verarbeitet.

10. Abschließende Grundsatzüberlegungen

Abschließend sollen einige grundsätzliche Überlegungen zur Bedeutung von diagnostischer Kompetenz bezogen auf das Fähigkeitsselbstkonzept erörtert werden. Dabei wird zum einen auf die subjektive Bedeutsamkeit des Fähigkeitsselbstkonzepts aus Sicht der Lehrkräfte eingegangen, zum anderen auf empirische Wirkungen diagnostischer Kompetenz.

10.1 Subjektive Bedeutsamkeit des Fähigkeitsselbstkonzepts

Unter Bezug auf das soziale Informationsverarbeitungsparadigma kann davon ausgegangen werden, dass für einen Abruf von Informationen (in unserem Fall der Einschätzung von Fähigkeitsselbstkonzepten) die betreffenden Informationen zunächst einmal enkodiert und gespeichert werden müssen. Die zentralen Prozesse des ersten Schrittes – der Enkodierung – sind Wahrnehmung und Aufmerksamkeitsfokussierung. Dabei werden erstens insbesondere solche Informationen wahrgenommen und fokussiert, zu denen Zugang besteht und zweitens solche, die als hoch relevant wahrgenommen werden (zsf. Abele & Gendolla, 1997). Ersteres wurde oben schon ausführlich diskutiert. In Bezug auf letzteres lässt sich für die vorliegende Thematik schlussfolgern: Schätzen Lehrkräfte die Relevanz des Fähigkeitsselbstkonzepts für den Unterricht und die Persönlichkeits- und Lernentwicklung der Schüler als eher gering ein, widmen sie diesem keine Aufmerksamkeit und können es in der Folge auch nicht adäquat einschätzen. Während in der Forschung die hohe Bedeutung von Selbstkonzepten für das Erleben und Verhalten von Personen unumstritten ist, weiß man bislang nahezu nichts darüber, welche Bedeutung Lehrkräfte diesem Schülermerkmal beimessen. Dabei ist erstens anzunehmen, dass Lehrkräfte, die dem Fähigkeitsselbstkonzept für den Unterricht und die Lernentwicklung der Schüler eine hohe Bedeutung zusprechen, tendenziell eher über eine hohe diagnostische Kompetenz verfügen und zweitens die Vermittlung von Diagnose- und Interventionsmöglichkeiten auch nur in diesem Fall sinnvoll ist.

10.2 Auswirkungen von diagnostischer Kompetenz

Ein grundlegendes Problem bisheriger Forschungen zu diagnostischer Kompetenz liegt darin, dass positive Auswirkungen diagnostischer Kompetenz zwar theoretisch plausibel abgeleitet werden können, empirisch jedoch bislang kaum erforscht wurden. Die einzigen deutschsprachigen Untersuchungen hierzu beziehen sich, wie bereits erwähnt, auf den Leistungsbereich (vgl. Hosenfeld, 2009; Schrader, 1989). Zu den potentiellen positiven Auswirkungen, die eine hohe diagnostische Kompetenz im Bereich des Fähigkeitsselbstkonzepts hat, gibt es hingegen bislang keinerlei Untersuchungen.

Zudem ist zum aktuellen Zeitpunkt ungeklärt, welches Effektivitätskriterium für diagnostische Kompetenz bezogen auf das Fähigkeitsselbstkonzept angemessen wäre. Für den Leistungsbereich ist relativ eindeutig, dass der Lernerfolg der Schüler ein geeignetes Kriterium darstellt (vgl. Helmke & Schrader, 1987; Schrader, 1989). Welches Kriterium jedoch eignet sich für den Bereich des Fähigkeitsselbstkonzepts? Ebenfalls der Lernerfolg der Schüler? Die Adaptivität von Unterricht (vgl. die Forderung für den Leistungsbereich von Weinert, Schrader & Helmke, 1990)? Eine realistische Selbstkonzeptentwicklung (vgl. z. B. Einsiedler, 1988; Fend, 1997)? Oder aber eine positive Selbstkonzeptentwicklung (s. z. B. Kammermeyer & Martschinke, 2004)?

Unabhängig davon, welches Effektivitätskriterium nun herangezogen wird, muss aus Sicht einer outputorientierten Sichtweise auf Schule kritisiert werden, dass sich bislang in diesem Bereich bezogen auf die Bedeutung von diagnostischer Kompetenz lediglich auf Plausibilitätsannahmen gestützt wird. Verfolgt man die Forderung nach messbaren Wirkungen von Unterricht und Schule (vgl. z. B. Terhart, 2002) konsequent, macht es jedoch erst dann Sinn, potentielle Einflussfaktoren und Möglichkeiten einer Entwicklung von diagnostischer Kompetenz zu untersuchen, wenn die Wirksamkeit von diagnostischer Kompetenz auch empirisch nachgewiesen ist. Hier besteht demnach dringender Forschungsbedarf.

Literatur

Abele, A.E. & Gendolla, G. (1997). Soziale Informationsverarbeitung. In J. Straub, W. Kempf & H. Werbik (Hrsg.), *Psychologie. Eine Einführung. Grundlagen, Methoden, Perspektiven* (S. 579–603). München: dtv.

Artelt, C. & Gräsel, C. (2009). Diagnostische Kompetenz von Lehrkräften. *Zeitschrift für Pädagogische Psychologie, 23*, 157–160.

Bayerisches Staatsministerium für Unterricht und Kultus (2005). *Lehrplan für die bayerische Grundschule* (5. Aufl.). München: Maiss.

Blatchford, P. (2009). Class size. In E. Anderman (Hrsg.), *Psychology of Classroom Learning: An Encyclopedia*. Detroit, MI: Macmillan Reference USA.

Bouffard, T., Marcoux, M.F., Vezeau, C. & Bordeleau, L. (2003). Changes in self-perceptions of competence and instrinsic motivation among elementary schoolchildren. *British Journal of Educational Psychology, 73,* 171–186.

Bos, W., Bonsen, M., Baumert, J., Prenzel, M., Selter, C. & Walther, G. (Hrsg.). (2008). *TIMSS 2007. Mathematische und naturwissenschaftliche Kompetenz von Grundschulkindern in Deutschland im internationalen Vergleich*. Münster: Waxmann.

Brühwiler, C. & Blatchford, P. (in Druck). Effects of class size and adaptive teaching competency on classroom processes and academic outcome. *Learning and Instruction*.

Chapman, J.W. & Tunmer, W.E. (1997). A longitudinal study of beginning reading achievement and reading self-concept: Self perception and performance. *British Journal of Educational Psychology, 67,* 279–291.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*, 155–159.

Cooksey, R.W. (1996). The methodology of social judgement theory. *Thinking and Reasoning*, *2*, 141–173.

Cooksey, R.W., Freebody, P. & Davidson, R. (1986). Teachers' predictions of children's early reading achievement: An application of social judgment theory. *American Educational Research Journal*, *23*, 41–64.

Cronbach, L.J. (1955). Processes affecting scores on "understanding of others" and "assumed similarity". *Psychological Bulletin*, *52,* 177–193.

Dickhäuser, O. (2006). Fähigkeitsselbstkonzepte: Entstehung, Auswirkung, Förderung. *Zeitschrift für Pädagogische Psychologie*, *20,* 5–8.

Dickhäuser, O. (2005). Teachers' inferences about students' self concepts – the role of dimensional comparison. *Learning and Instruction*, *15*, 226–235.

Dresel, M. (2001). Reattributionstraining. In W.H. Honal (Hrsg.), *Handbuch der Schulberatung* (S. 1–18). Landsberg: mvg.

Dresel, M. & Ziegler, A. (2006). Langfristige Förderung von Fähigkeitsselbstkonzept und impliziter Fähigkeitstheorie durch computerbasiertes attributionales Feedback. *Zeitschrift für Pädagogische Psychologie*, *20*, 49–63.

Einsiedler, W. (1988). Schulanfang und Persönlichkeitsentwicklung. *Grundschule*, *20*, 20–23.

Einsiedler, W., Martschinke, S. & Kammermeyer, G. (2007). *Die Grundschule zwischen Heterogenität und gemeinsamer Bildung. Berichte und Arbeiten aus dem Institut für Grundschulforschung, Nr. 105.* Universität Erlangen-Nürnberg.

Federer, M., Stüber, S., Margraf, J., Schneider, S. & Herrle, J. (2001). Selbst- und Fremdeinschätzungen der Kinderängstlichkeit. *Zeitschrift für Differentielle und Diagnostische Psychologie*, *22*, 194–205.

Fend, H. & Stöckli, G. (1997). Der Einfluß des Bildungssystems auf die Humanentwicklung: Entwicklungspsychologie der Schulzeit. In F. E. Weinert (Hrsg.), *Psychologie des Unterrichts und der Schule* (S. 1–35). Göttingen: Hogrefe.

Fuchs-Heinritz, W. (2007). Begrenzte Rationalität. In W. Fuchs-Heinritz, R. Lautmann, O. Rammstedt & H. Wienold (Hrsg.), *Lexikon zur Soziologie* (4., überarb. Aufl., S. 533). Wiesbaden: VS Verlag für Sozialwissenschaften.

Gabriel, K., Mösko, E. & Lipowsky, F. (in Druck). Selbstkonzeptentwicklung von Jungen und Mädchen – Ergebnisse aus der PERLE-Studie. In F. Günther & F. Hellmich (Hrsg.), *Selbstkonzepte im Grundschulalter*. Stuttgart: Kohlhammer.

Gigerenzer, G. & Selten, R. (2001). *Bounded rationality: The adaptive toolbox*. Cambridge, MA: MIT Press.

Greb, K. (2009). *Kompetenzmodellierung des diagnostischen Urteils bei Grundschullehrern*. Unveröffentlichte Dissertation, Universität Kassel.

Greb, K., Lipowsky, F., & Faust, G. (in Vorbereitung). Instrumente (MZP2 & 3): Schüler und Eltern. In F. Lipowsky, G. Faust & K. Greb (Hrsg.), *Dokumentation der Erhebungsinstrumente des Projekts „Persönlichkeits- und Lernentwicklung von Grundschulkindern" (PERLE). Materialien zur Bildungsforschung*. Frankfurt a.M.: GFPF.

Harter, S. (2006). The self. In N. Eisenberg (Hrsg.), *Handbook of child psychology: Vol. 3. Social, emotional, and personality development* (6. Auflage, S. 505–570). New York: Wiley.

Helmke, A. (2009). *Unterrichtsqualität und Lehrerprofessionalität. Diagnose, Evaluation und Verbesserung des Unterrichts*. Seelze: Klett-Kallmeyer.

Helmke, A. (1998). Vom Optimisten zum Realisten? Zur Entwicklung des Fähigkeitsselbstkonzepts vom Kindergarten bis zur sechsten Klassenstufe. In F. E. Weinert (Hrsg.), *Entwicklung im Kindesalter* (S. 115–132). Weinheim: Psychologie Verlags Union.

Helmke, A. & Fend, H. (1981). Wie gut kennen Eltern ihre Kinder und Lehrer ihre Schüler? Ergebnisse der Konstanzer Untersuchungen zu Bedingungen der Diagnosegenauigkeit bei Eltern und Lehrern. In G. Zimmer (Hrsg.), *Persönlichkeitsentwicklung und Gesundheit im Schulalter* (S. 341–360). Frankfurt a.M.: Campus.

Helmke, A., Hosenfeld, I. & Schrader, F.-W. (2004). Vergleichsarbeiten als Instrument zur Verbesserung der Diagnosekompetenz von Lehrkräften. In R. Arnold & C. Griese (Hrsg.), *Schulleitung und Schulentwicklung* (S. 119–144). Hohengehren: Schneider.

Helmke, A. & Schrader, F.-W. (1987). Interactional effects of instructional quality and teacher judgment accuracy on achievement. *Teaching and Teacher Education*, *3*, 91–98.

Hoge, R.D., & Colardarci, T. (1989). Teacher based judgments of academic achievement: A review of the literature. *Review of Educational Research*, *59*, 297–313.

Hosenfeld, I. (2009, August). *Teachers accuracy in judging mathematics tasks and the relationship to achievement gains*. Paper presented at the 13th biennial conference EARLI 2009, Amsterdam.

Ingenkamp, K., Petillon, H. & Weiß, M. (Hrsg.). (1985). *Klassengröße: Je kleiner, desto besser? Forschungs- und Diskussionsstand zur Wirkung der Klassenfrequenz*. Weinheim: Beltz.

Kammermeyer, G. & Martschinke, S. (2004). KILIA – Selbstkonzept- und Leistungsentwicklung im Anfangsunterricht. In G. Faust, M. Götz, H. Hacker & H.-G. Rossbach (Hrsg.), *Anschlussfähige Bildungsprozesse im Elementar- und Primarbereich* (S. 204–217). Bad Heilbrunn: Klinkhardt.

Kammermeyer, G. & Martschinke, S. (2006). Selbstkonzept- und Leistungsentwicklung in der Grundschule – Ergebnisse aus der KILIA-Studie. *Empirische Pädagogik*, *20*, 245–259.

Karing, C. (2009). Diagnostische Kompetenz von Grundschul- und Gymnasiallehrkräften im Leistungsbereich und im Bereich Interessen. *Zeitschrift für Pädagogische Psychologie*, *23*, 197–209.

Krolak-Schwerdt, S., Böhmer, M. & Gräsel, C. (2009). Verarbeitung von schülerbezogener Information als zielgeleiteter Prozess: Der Lehrer als „flexibler Denker“. *Zeitschrift für Pädagogische Psychologie*, *23*, 175–186.

Krolak-Schwerdt, S. & Rummer, R. (2005). Der Einfluss von Expertise auf den Prozess der schulischen Leistungsbeurteilung. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, *37*, 205–213.

Kultusministerkonferenz (2004). *Standards für die Lehrerbildung: Bildungswissenschaften*. Zugriff am 15.10.2009 unter http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2004/2004_12_16-Standards-Lehrerbildung.pdf

Lipowsky, F., Faust, G. & Greb, K. (Hrsg.) (2009). *Dokumentation der Erhebungsinstrumente des Projekts „Persönlichkeits- und Lernentwicklung von Grundschülern“ (PERLE) – Teil 1*. Frankfurt a.M.: GFPF. (Materialien zur Bildungsforschung; Bd. 23).

Lorenz, C. & Artelt, C. (2009). Fachspezifität und Stabilität diagnostischer Kompetenz von Grundschullehrkräften in den Fächern Deutsch und Mathematik. *Zeitschrift für Pädagogische Psychologie, 23,* 211–222.

Marsh, H.W. (1984). Relations among dimensions of self-attribution, dimensions of self-concept and academic achievements. *Journal of Educational Psychology*, *76*, 1291–1308.

Marsh, H.W. (1990). A multidimensional, hierarchical model of self-concept: Theoretical and empirical justification. *Educational Psychology Review*, *2*, 77–172.

Marsh, H.W., Smith, I.D. & Barnes, J. (1984). Multidimensional self-concepts: Relationships with inferred self-concepts and academic achievement. *Australian Journal of Psychology*, *30*, 367–386.

Möller, J., & Trautwein, U. (2009). Selbstkonzept. In E. Wild & J. Möller (Hrsg.), *Pädagogische Psychologie* (S. 179–203). Heidelberg: Springer.

Moschner, B. & Dickhäuser, O. (2006). Selbstkonzept. In D.H. Rost (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (3. Aufl., S. 685–692). Weinheim: Beltz PVU.

Muthén, L.K. & Muthén, B.O. (1998-2007). *Mplus User's Guide* (5th ed.). Los Angeles, CA: Muthén & Muthén.

Padberg, F. (2007). *Didaktik der Arithmetik für Lehrerausbildung und Lehrerfortbildung* (3. Aufl.) Berlin: Springer.

Pitman, E.J.G. (1939). A note on normal correlation. *Biometrika*, *31*, 9–12.

Pohlmann, B., Möller, J. & Streblow, L. (2004). Zur Fremdeinschätzung von Schülerselbstkonzepten durch Lehrer und Mitschüler. *Zeitschrift für Pädagogische Psychologie*, *18*, 157–169.

Poloczek, S. (2007). *Zur Struktur des Fähigkeitsselbstkonzepts von Erstklässlerinnen und Erstklässlern*. Unveröffentlichte Diplomarbeit, Johann-Wolfgang-Goethe-Universität Frankfurt a.M.

Poloczek, S., Greb, K. & Lipowsky, F. (2010). *Facetten des schulischen Selbstkonzepts von Schulanfängern*. Manuskript eingereicht zur Publikation.

Praetorius, A.-K., Greb, K., Dickhäuser, O. & Lipowsky, F. (in Druck). Wie gut schätzen Lehrer die Fähigkeitsselbstkonzepte ihrer Schüler ein? Zur diagnostischen Kompetenz von Lehrkräften. *Psychologie in Erziehung und Unterricht*.

Raudenbush, S.W., Bryk, A., Cheong, Y.F. & Congdon, R. (2004). *HLM 6: Hierarchical linear and nonlinear modeling*. Chicago, IL: Scientific Software International.

Schilling, S.R., Sparfeldt, J.R. & Rost, D.H. (2006). Facetten schulischen Selbstkonzepts. Welchen Unterschied macht das Geschlecht? *Zeitschrift für Pädagogische Psychologie*, *20*, 9–18.

Schöne, C., Dickhäuser, O., Spinath, B. & Stiensmeier-Pelster, J. (2002). *Skalen der Erfassung des schulischen Selbstkonzepts – SESSKO*. Göttingen: Hogrefe.

Schrader, F.-W. (1989). *Diagnostische Kompetenzen von Lehrern und ihre Bedeutung für die Gestaltung und Effektivität des Unterrichts*. Frankfurt a.M.: Peter Lang.

Schrader, F.-W. (2009). Anmerkungen zum Themenschwerpunkt Diagnostische Kompetenz von Lehrkräften. *Zeitschrift für Pädagogische Psychologie*, *23*, 237–245.

Schrader, F.-W. (in Druck). Lehrer als Diagnostiker. In E. Terhart, H. Bennewitz & M. Rothland (Hrsg.), *Handbuch der Forschung zum Lehrerberuf*. Münster: Waxmann.

Schrader, F.-W. & Helmke, A. (1987). Diagnostische Kompetenz von Lehrern: Komponenten und Wirkungen. *Empirische Pädagogik*, *1*, 27–52.

Schrader, F.-W. & Helmke, A. (1990). Lassen sich Lehrer bei der Leistungsbeurteilung von sachfremden Gesichtspunkten leiten? Eine Untersuchung zu Determinanten diagnostischer Lehrerurteile. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, *22*, 312–324.

Schrader, F.-W., Helmke, A., Hosenfeld, I., Halt, A.C. & Hochweber, J. (2006). Komponenten der Diagnosegenauigkeit von Lehrkräften: Ergebnisse aus Ver-

gleichsarbeiten. In F. Eder, A. Gastager & F. Hofmann (Hrsg.), *Qualität durch Standards? Beiträge zur 67. AEPF-Tagung in Salzburg* (S. 265–278). Münster: Waxmann.

Snijders, T.A.B. & Bosker, R.J. (1999). *Multilevel analysis. An introduction to basic and advanced multilevel modeling*. London: Sage.

Spinath, B. (2004). Determinanten von Fähigkeitsselbstwahrnehmungen im Grundschulalter. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, *36*, 63–68.

Spinath, B. (2005). Akkuratheit der Einschätzung von Schülermerkmalen durch Lehrer und das Konstrukt der diagnostischen Kompetenz. *Zeitschrift für Pädagogische Psychologie*, *19*, 85–95.

Südkamp, A. & Möller, J. (2009) Referenzgruppeneffekte im Simulierten Klassenraum: direkte und indirekte Einschätzungen von Schülerleistungen. *Zeitschrift für Pädagogische Psychologie, 23,* 161–174.

Terhart, E. (2002). *Standards für die Lehrerbildung. Eine Expertise für die Kultusministerkonferenz.* Universität Münster: Zentrale Koordination Lehrerbildung.

Trautwein, U. (2003). *Schule und Selbstwert. Entwicklungsverlauf, Bedeutung von Kontextfaktoren und Effekte auf die Verhaltensebene*. Münster: Waxmann.

Weinert, F.E. & Schrader, F.-W. (1986). Diagnose des Lehrers als Diagnostiker. In H. Petillon, J. Wagner & B. Wolf (Hrsg.), *Schülergerechte Diagnose. Theoretische und empirische Beiträge zur Pädagogischen Diagnostik* (S. 11–29). Weinheim: Beltz.

Weinert, F.E., Schrader, F.-W. & Helmke, A. (1990). Educational expertise: Closing the gap between educational research and classroom practice. *School Psychology International*, *11*, 163–180.

Wiater, W. (2006). *Theorie der Schule* (2. Aufl.). Donauwörth: Auer.

Wild, K-P. (1991). *Identifikation hochbegabter Schüler: Lehrer und Schüler als Datenquellen*. Heidelberg: Asanger.

Wild, K-P. & Rost, D. H. (1995). Klassengröße und Genauigkeit von Schülerbeurteilungen. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, *27*, 79–90.

Whitehead, J.M. (1996). Sex stereotypes, gender identity and subject choice at A-level. *Educational Research*, *38*, 147–160.

Ziegler, A. & Schober, B. (2001). *Theoretische Grundlagen und praktische Anwendung von Reattributionstrainings*. Regensburg: Roderer.

Anna-Katharina Praetorius, M.A., Promotionsstipendiatin im DFG-Graduiertenkolleg Unterrichtsprozesse, Universität Koblenz-Landau
E-Mail: praetorius@uni-landau.de

Karina Greb, Dr. des., Wissenschaftliche Mitarbeiterin im Projekt PERLE, Universität Kassel und Deutsches Institut für Internationale Pädagogische Forschung (DIPF), Frankfurt am Main
E-Mail: greb@dipf.de

Frank Lipowsky, Prof. Dr., Leiter des Fachgebiets Empirische Schulforschung, Universität Kassel
E-Mail: lipowsky@uni-kassel.de

Mario Gollwitzer, Prof. Dr., Professur Psychologische Methodenlehre, Philipps-Universität Marburg
E-Mail: mario.gollwitzer@staff.uni-marburg.de