



Diagnostik in der Lehrerbildung

Vortrag zum 43. BAK Seminartag in Saarbrücken
02.10.2009

Univ. Prof. Dr. Roland Brünken

Lehrstuhl für Empirische Bildungsforschung
FR Erziehungswissenschaft - Universität des Saarlandes

Überblick



- Einführung
 - PISA und die diagnostische Kompetenz von Lehrern
 - Definitionsversuche
 - Funktionen schulischer Diagnostik
- Methodische Grundlagen
 - Arten und Gegenstandsbereiche von Diagnostik
 - Gütekriterien
 - Komponenten der Urteilsgenauigkeit
 - Indikatoren der Urteilsgenauigkeit
- Empirische Befunde zur diagnostischen Kompetenz von Lehrern
 - Beurteilung von Schülerleistung
 - Diagnostische Kompetenz und Unterrichtserfolg
 - Beurteilung von Lernvoraussetzungen
 - Konstrukthomogenität
 - Förderung diagnostischer Kompetenzen
- Diagnostik in der universitären Lehramtsausbildung
- Forschungsbedarf
- Zusammenfassung und Ausblick



PISA und die diagnostische Kompetenz von Lehrern

- PISA 2001: Eine Teilstichprobe von Hauptschullehrer/innen wurde aufgefordert Schüler/innen zu benennen, die aufgrund geringer Lesefähigkeit ernsthafte Probleme beim Übergang in das Berufsleben haben würden. (Delegationsverfahren)
- Dabei zeigte sich, dass von den Lehrer/innen deutlich weniger Schüler/innen benannt wurden, als laut der Testergebnisse der so definierten Kompetenzstufe zuzuordnen waren.
- Wie gerechtfertigt ist der Schluss auf geringe diagnostische Kompetenzen von Lehrer/innen und was sind überhaupt diagnostische Kompetenzen?

		PISA 2000 Lese-Kompetenzstufe		
		< K1	K1	> K1
Lehrerdiagnose	Schwacher Leser	11,4 %	3,7%	2,8%
	Kein schwacher Leser	88,6 %	96,3 %	97,2 %



PISA und die diagnostische Kompetenz von Lehrern (2)

- Es wurden nur Hauptschullehrer/innen untersucht. Dies ist **keine repräsentative Stichprobe** für alle Lehrer/innen.
- Es wurde nur **ein** zu diagnostizierendes **Schülermerkmal** untersucht, nämlich Leseleistung.
- Es wurde nur **ein Akkuratheitsmaß** herangezogen (ein Abweichungsmaß zwischen Lehrerurteil und Schülermerkmalsausprägung).
- Die **Beurteilung der vorgefundenen Werte** als „unzureichend“ ist **willkürlich**, da es keinen verbindlichen Maßstab gibt.
- **Andere** vorliegende **Forschungsergebnisse**, die diesem Befund **widersprechen**, werden außer acht gelassen.

Aus diesem Teilergebnis von PISA kann nicht allgemein auf geringe diagnostische Kompetenzen von Lehrer/innen geschlossen werden. (Spinath, 2004)

Aber die Frage bleibt: Was ist diagnostische Kompetenz und wie steht es um deren Ausprägung bei Lehrerinnen und Lehrern?



Definitionsversuche

- **KMK Richtlinien** (Standards für die Lehrerbildung, KMK 2004)
 - 11 Kompetenzen – zwei mit direktem diagnostischen Bezug
 - Kompetenz 7:
Lehrerinnen und Lehrer diagnostizieren Lernvoraussetzungen und Lernprozesse von Schülerinnen und Schülern, ...
 - Kompetenz 8:
Lehrerinnen und Lehrer erfassen Leistungen von Schülerinnen und Schülern auf der Grundlage transparenter Beurteilungsmaßstäbe



Definitionsversuche (2)

- **Bildungswissenschaften:**
 - Das Konstrukt der **diagnostischen Kompetenz** bezeichnet die Fähigkeit, Merkmale von Personen korrekt einzuschätzen. (nach Schrader, 2001)
 - Diagnosen stellen in der Regel **explizite** Aussagen über Zustände und Merkmale von Personen dar, die Ergebnis eines **reflektierten** und **methodisch kontrollierte** durchlaufenen diagnostischen Prozesses sind
 - Diagnostische Urteile dienen dabei in der Regel der **Entscheidungsfindung**



Funktionen schulischer Diagnostik

- **Mikroebene:**
 - Unterrichtsplanung
 - (Vergleichende) Analyse von Gruppenleistungen zur Anpassung von Unterrichtsparametern (Stoffmenge, Geschwindigkeit etc.)
 - Förderdiagnostik
 - Diagnose individueller (Leistungs-) Entwicklung zur Auswahl angemessener Fördermaßnahmen (Binnendifferenzierung)
- **Makroebene:**
 - Selektionsdiagnostik
 - Zur Vorbereitung schulrelevanter Entscheidungsprozesse (z.B. Übergangentscheidungen)
 - Systemdiagnostik
 - Zur (vergleichenden) Analyse der Ergebnisse von Bildungsprozessen und –systemen (Output-Diagnostik; z.B. PISA; Lernstandserhebungen)

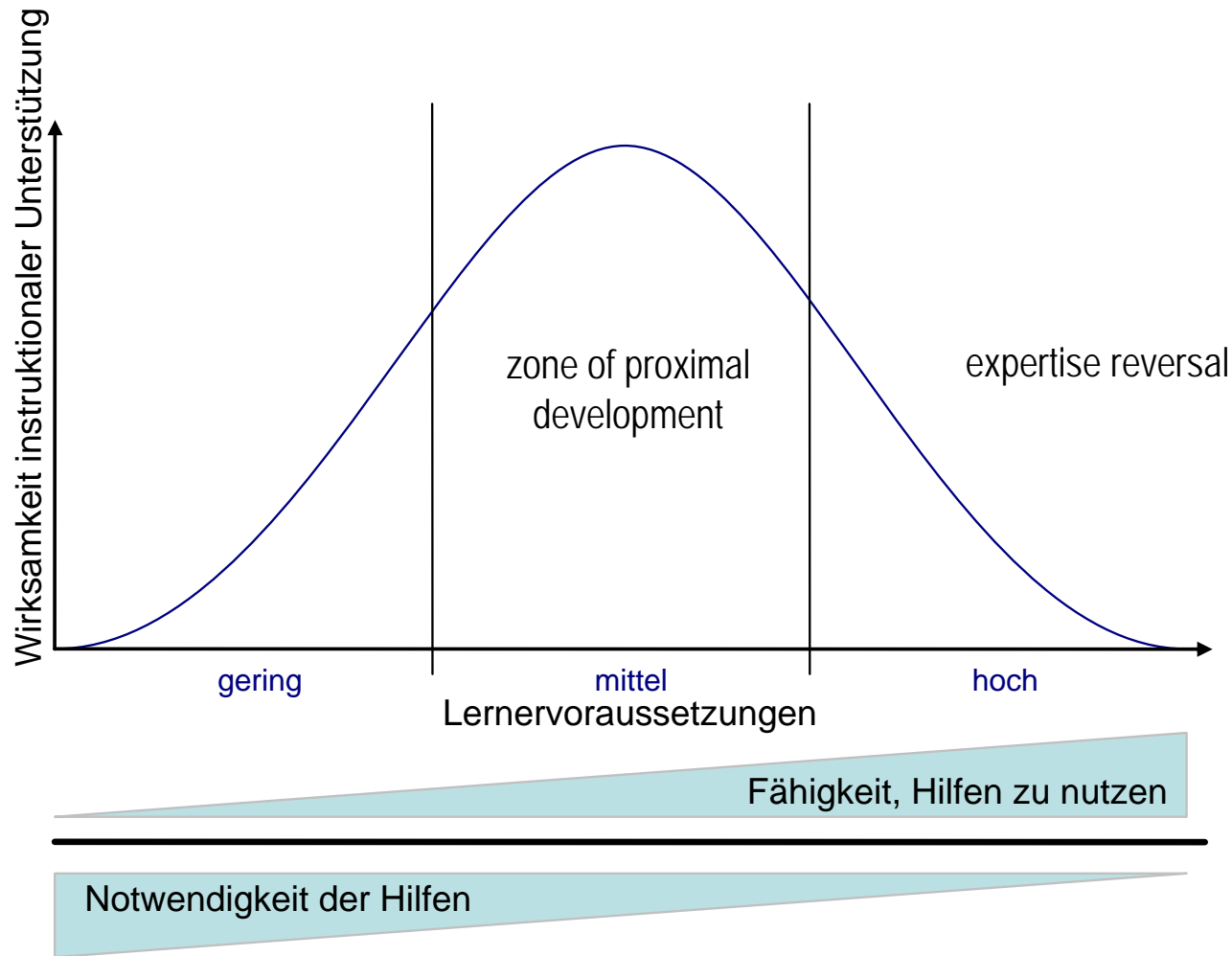


Funktionen schulischer Diagnostik (2)

- Warum überhaupt Diagnostik?
- Idee: „Erziehungsmaßnahmen erzielen eine optimale Wirkung, wenn eine hinreichende Passung zwischen den gestellten Anforderungen und den Persönlichkeitsmerkmalen der Betroffenen vorliegt (Schrader, 2001)
- ATI: Aptitude Treatment Interaction (Brünken & Leutner, 2005)

Funktionen schulischer Diagnostik (3):

ATI - Beispiel: Expertise Reversal Effekt (Kalyuga et al, 2003)



(nach Seufert, Jänen & Brünken, 2007)



Arten von Diagnostik

- Standardisierte pädagogisch-psychologische Diagnostik (nicht Aufgabe von Lehrern)
 - Psychometrische Intelligenzdiagnostik (IQ)
 - Diagnostik von Teilleistungsstörungen (z.B. LRS)
 - Klinische Verhaltensdiagnostik (z.B. ADHS)
- Nicht standardisierte, informelle Diagnostik
 - Leistungsbeurteilung (schriftlich/mündlich)
 - Verhaltensdiagnostik (Sozialverhalten, Arbeitsstil etc..)
 - Motivations- und Volitionsdiagnostik
 - Verdachtsdiagnostik von Leistungsauffälligkeiten (sowohl im oberen wie im unteren Leistungsbereich)
 - Diagnostik von Verhaltensauffälligkeiten (z.B. Unterrichtsstörungen, Mobbing etc.)



Gütekriterien

- **Validität:** Gültigkeit (z.B. Leistung oder Anstrengung)
- **Reliabilität:** Zuverlässigkeit (z.B. kommen zwei Urteile zum gleichen Ergebnis?)
- **Objektivität:** Unabhängigkeit vom Beurteiler/Auswerter (z.B. käme ein anderer Urteiler zum gleichen Ergebnis?)
- In wieweit gelten die Kriterien standardisierter Diagnostik auch für Lehrerurteile?
- **Veridikalität:** wie gut stimmt ein Urteil mit dem tatsächlichen (objektiv gemessenen) Merkmal überein (Helmke, et al. 2004)



Komponenten der Urteilsgenauigkeit (Schrader, 1989)

- Wissen über relevante diagnostische Merkmale (diagnostisches Wissen)
- Urteilsmaßstäbe (Bezugsnormen)
 - Individuell
(Vergleich des Schülers mit sich selbst – zur individuellen Veränderungsmessung)
 - Sozial
(Vergleich innerhalb einer Gruppe (z.B. Klasse) – z.B. zur Wettbewerbsförderung)
 - Kriterial
(Vergleich mit einem (extern) vorgegeben Kriterium – z.B. zur Kompetenzmessung oder zum externen Vergleich)
- Urteilstendenzen
 - Halo-Effekte (Überstrahlungseffekt)
 - Milde/Strenge-Effekte
 - Primacy/Recency-Effekt (Reihenfolge-Effekte)



Indikatoren der Urteilsgenauigkeit

- **Generelle Akkuratheitsmaße:**
 - Übereinstimmung mit / Abweichung von einem Außenkriterium (Standards, Selbsteinschätzung, Tests)
 - Vorhersagegenauigkeit
- **Komponenten des Lehrerurteils (Schrader & Helmke, 1987; Helmke et al, 2004)**
 - **Niveaue**komponente (mittlere Merkmalsausprägung in der Klasse)
 - **Differenzierungs**komponente (Merkmalshomogenität in der Klasse)
 - **Vergleichs**komponente (Rangordnung der Schüler untereinander)

Empirische Befunde zur diagnostischen Kompetenz von Lehrern



- In der Regel werden in empirischen Studien diagnostische Kompetenzen anhand von Akkuratheitsmaßen beurteilt
- Methodisches Vorgehen:
 - Schüler werden Aufgaben zur Lösung vorgelegt – die Lehrer sollen einschätzen, ob/in wieweit Schüler Erfolg haben
 - Vergleich Lehrerurteil/Schülerleistung (Korrelationen)
 - Vergleich Lehrerurteil mit standardisierten Test- und Fragebogenverfahren (z.B. IQ Tests; Fragebögen zum FSK; DISK-Gitter etc)
- Überwiegende betrachtete Lehrerurteile
 - Schülerleistung
 - Leistungsvoraussetzungen
 - Intelligenz
 - Lern- und Leistungsmotivation
 - (akademisches) Fähigkeitsselbstkonzept
 - Leistungsängstlichkeit



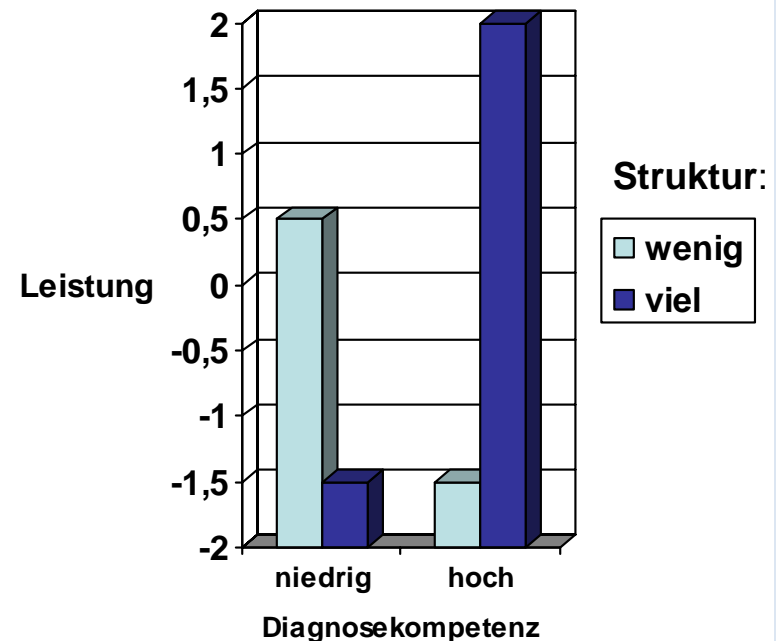
Beurteilung von Schülerleistungen

- Lehrer neigen zu einer (leichten) **Niveauüberschätzung** (Schrader & Helmke, 1987)
 - ggf. getragen von normativen Aspekten (was müssten Schüler können?)
 - eher kompetenz- als leistungsorientiert (was könnten die Schüler *eigentlich* leisten?)
- Lehrer neigen dazu, Unterschiede zu akzentuieren, d.h. die **Differenzierungs**komponente wird eher überschätzt
- Die größte Heterogenität zwischen Lehrern zeigt sich in Bezug auf die **Vergleichs**komponente (Rangordnung der Schüler untereinander)
 - das Spektrum der Übereinstimmung mit einem Außenkriterium (Rangordnung anhand der Leistung in einem Test) reicht von $r = .04$ bis $r = .88$
 - bei im Durchschnitt mittlerer Übereinstimmung ($r = .64$; Helmke et al, 2004)
- Durchschnittliche mittlere Übereinstimmung zwischen Leistungstest und Lehrerurteil $r = .66$ (Metaanalyse von Hoge & Coladarci, 1989)



Diagnostische Kompetenz und Unterrichtserfolg

- Zusammenhang Diagnostische Kompetenz und der Wirkung von Strukturierung und Unterstützung (Weinert & Helmke, 1987):
 - „...hohe diagnostische Kompetenz erleichtert es den Lehrern, lernförderliche unterstützende Maßnahmen zielgerichtet einzusetzen
 - bei geringer diagnostischer Kompetenz erweisen sich Unterstützungsmaßnahmen als ineffizient (Fehlallokation von Förderung)



nach Weinert & Helmke, 1987



Beurteilung von Lernvoraussetzungen

- **Intelligenz**
 - untersucht im Rahmen der Diagnostik von Hochbegabung (Marburger Hochbegabtenprojekt; Wild, 1992)
 - Gute mittlere Übereinstimmung zwischen Lehrereinschätzung und allgemeiner Intelligenz ($r = .67$)
 - Übereinstimmung bei der Einschätzung spezifischer Faktoren (verbal, numerisch, räumlich) geringer
 - Aber: Lehrereinschätzung hoch leistungskorreliert, d.h. IQ wird aus hoher Leistung geschlossen
 - Daher: Probleme bei der Identifikation von Underachievern (hoher IQ – niedrige Leistung)



Beurteilung von Lernvoraussetzungen (2)

- Wenig Studien zu nicht kognitiven Lernvoraussetzungen
- Akkuratheitsmaß: Korrelation Lehrerurteil / Schülerelbsteinschätzung
- Ausgewählte Ergebnisse:
 - **Fähigkeitsselbstkonzept**: mittlere Zusammenhänge ($.30 < r < .60$; Marsch, 1990)
 - **Lernmotivation**: geringe positive Zusammenhänge ($r < .30$; Helmke & Fend, 1981)
 - **Leistungsängstlichkeit**: schwache positive Zusammenhänge ($r < .30$; Böhnke et al, 1986)
- Wiederum Hinweise auf den Einfluss von Leistungen auf die Einschätzung (z.B. gute Schüler werden als motivierter angesehen als weniger gute)



Konstruktthomogenität

- B. Spinath, 2005: Akkuratheit von Lehrerurteilen (ZfPP)
- Fragestellung: gibt es eine generelle Beurteilungsfähigkeit im Sinne diagnostischer Kompetenz?
- Voraussetzung: Übereinstimmende Beurteilungsqualität (Akkuratheit) über unterschiedliche Bereiche und verschiedene Komponenten
- Die Ergebnisse zeigen nur geringe Zusammenhänge (Korrelationen) zwischen den verschiedenen Komponenten und Bereichen
- Fazit: bereichsspezifisch hohe Akkuratheitsvarianz ibs in der Vergleichskomponente und in nicht-kognitiven Bereichen
- Folgerung: es gibt nicht „die diagnostische Kompetenz“



Förderung diagnostischer Kompetenzen

- Kaum Empirie!
- Keine längsschnittlichen Studien zur Kompetenzentwicklung
- Pragmatische Lösungsvorschläge
 - Schaffung von Übungs- und Reflexionsmöglichkeiten (Wahl, Weinert & Huber, 1997)
 - Explikation relevanter Beobachtungsmerkmale (proximale Merkmale) für nicht beobachtbare (latente, distale) Konstrukte (Linsenmodell; Kleber 1992)
 - Nutzung schulübergreifender Kompetenzerhebungen und ihrer Leistungsrückmeldungen zur vergleichenden Einschätzung der (eigenen) diagnostischen Kompetenz (z.B. in VERA; Helmke et al. 2004)

Zusammenfassung empirischer Forschungsergebnisse



- Leistungsdiagnostik funktioniert relativ gut bei
 - Tendenz zur Verwendung sozialer Bezugsnormen
 - Tendenz zur generellen Überschätzung der Schülerleistungen (Niveauelemente)
 - fehlenden externen Bezugskriterien
- Intelligenzdiagnostik relativ zuverlässig, aber
 - hoch leistungskorreliert
 - dadurch Probleme bei der Identifikation von Underachievern
- Nicht kognitive Schülermerkmale
 - Beurteilungsgüte umso besser, je enger das Merkmal mit Leistung korreliert ist
 - Unterschätzung von Leistungsängstlichkeit
- Ein globales Konstrukt „diagnostische Kompetenz“ lässt sich empirisch nicht zeigen (Spinath, 2004; 2005)
- Es besteht erheblicher Ausbildungsbedarf in Hinblick auf diagnostisches Wissen und diagnostischer Methodik

Diagnostik in der universitären Lehramtsausbildung an der UdS



- Das Saarbrücker Modell der Bildungswissenschaften in der Lehrerbildung (UdS SPO 2007)
- Modularisierter, an den KMK Richtlinien und der Empfehlungen von DGfE und DGPs orientierter Studiengang mit 48 CP (>25% der Lehramtsausbildung)
- 4 Kompetenzbereiche
 - K1: Lehren und Lernen
 - K2: Persönlichkeitsentwicklung und Erziehung
 - **K3: Diagnostik, Beratung, Intervention**
 - K4: Schulentwicklung, Qualitätssicherung & Systemmonitoring

Diagnostik in der universitären Lehramtsausbildung an der UdS



- Diagnostische Inhalte in K3
 - Einführung in die Diagnostik (VL)
 - Schulische Leistungsdiagnostik
 - Förderdiagnostik
 - Standardisierte Diagnostik (psychometrische Testverfahren)
- Diagnostik in den anderen Kompetenzbereichen
 - Bildungswissenschaftliche Grundlagen (K1)
 - Empirische Forschungsmethoden (K1)
 - Motivation & Motivationsdiagnostik (K1)
 - Diagnostik von Schulschwierigkeiten und Disziplinproblemen (K2)
 - (International) Vergleichende Bildungssystemdiagnostik (K4)



Forschungsbedarf

- Zur Kompetenzdiagnostik fehlen derzeit noch brauchbare, schulfachbezogene Kompetenzmodelle und darauf basierende Testverfahren
 - DFG Schwerpunktprogramm "Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen" (Klieme & Leutner, 2008)
- Individuelle Voraussetzung erfolgreicher Lehramtsstudierender sind noch unklar
 - BMBF Programm Professionalisierung der Lehrerbildung
 - z.B. Projekt SIOS-L (Brünken, Spinath & Bedersdorfer, 2009)
- Evaluationsstudien zur Effizienz von Lehreraus- und weiterbildung fehlen noch weitgehend



Zusammenfassung und Ausblick

- Generelle Kritik an der diagnostischen Kompetenz von Lehrern ist nicht gerechtfertigt
- Ibs. im Leistungsbereich sind Lehrer zuverlässige, wenn auch in der Regel „milde“ Diagnostiker
- Diagnostische Urteile von Lehrern haben in allen Bereichen einen „Leistungs-Bias“
- Die universitäre Ausbildung dient in erster Linie der Vermittlung von
 - diagnostischem Wissen
 - diagnostischen Methoden und Techniken
 - Grundlagen der Urteilsbildung und ihrer Schwierigkeiten
 - Grundlagen der standardisierten Diagnostik
- Das Konzept der diagnostischen Kompetenz als eindimensionales Konstrukt ist umstritten; insgesamt ist eine differenziertere Perspektive einzunehmen



Vielen Dank für Ihr Interesse

Univ. Prof. Dr. Roland Brünken

Lehrstuhl für Empirische Bildungsforschung
FR Erziehungswissenschaft - Universität des
Saarlandes



Literatur

- Brünken, R. & Leutner, D. (2005). Individuelle Unterschiede beim Lernen mit Neuen Medien. In: Schilling, S.R., Sparfeldt, J.R. & Pruisken, C.(Hrsg.) *Aktuelle Aspekte pädagogisch-psychologischer Forschung. Detlef H. Rost zum 60. Geburtstag* (S. 25-40). Münster: Waxmann.
- Böhnke, K., Silbereisen, R. K., Reynolds, C. R. & Richmond, B. O. (1986). What I think and feel: German experience with the revised form of the Children's Manifest Anxiety Scale. *Personality and Individual Differences*, 7, 553–560.
- Cronbach, L. J. (1957). Two disciplines of scientific psychology. *American Psychologist*, 12, 671-684.
- Helmke, A. (2004). *Unterrichtsqualität. Erfassen, Bewerten, Verbessern*. Seelze: Kallmeyer.
- Helmke, A.&Fend, H. (1981).Wie gut kennen Eltern ihre Kinder und Lehrer ihre Schüler? In G. Zimmer (Hrsg.), *Persönlichkeitsentwicklung und Gesundheit im Grundschulalter: Gefährdungen und Prävention* (S. 341–360). Frankfurt: Campus.
- Helmke, A., Hosenfeld, I. & Schrader, F.-W. (2004). Vergleichsarbeiten als Instrument zur Verbesserung der Diagnosekompetenz von Lehrkräften. In R. Arnold & C. Griese (Hrsg.), *Schulmanagement und Schulentwicklung*. Hohengehren: Schneider-Verlag.
- Hoge, R. D. & Coladarci, T. (1989). Teacher-based judgments of academic achievement:Areviewof literature. *Review of Educational Psychology*, 59, 297–313.
- Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). Expertise reversal effect. *Educational Psychologist*, 38, 23-31.



Literatur 2

- Kleber, E.W. (1992). *Diagnostik in pädagogischen Handlungsfeldern*. München: Juventa.
- Marsh, H.W. (1990). A multidimensional, hierarchical model of self-concept: Theoretical and empirical justification. *Educational Psychology Review*, 2, 77–172.
- Schrader, F.-W. (2001). Diagnostische Kompetenz von Eltern und Lehrern. In D. H. Rost, *Handwörterbuch Pädagogische Psychologie* (S. 91-96). Weinheim: Beltz.
- Schrader, F.W. & Helmke, A. (1989). Diagnostische Kompetenz von Lehrern: Komponenten und Wirkungen. *Empirische Pädagogik*, 1, 27-52.
- Seufert, T., Jänen, I. & Brünken R. (2007). The impact of intrinsic cognitive load on the effectiveness of graphical help for coherence formation. *Computers in Human Behavior*, 23, 1055-1071.
- Spinath, B. (2004). Diagnostische Kompetenzen von Lehrerinnen und Lehrern. *Lernende Schule*, 26, 16-17.
- Spinath, B. (2005). Akkuratheit der Einschätzung von Schülermerkmalen durch Lehrer/innen und das Konstrukt der diagnostischen Kompetenz. *Zeitschrift für Pädagogische Psychologie*, 19, 85-95.
- Wahl, D., Weinert, F.E. & Huber, G. (1997). *Psychologie für die Schulpraxis*. München: Kösel.
- Westhoff, K. & Kluck, M.-L. (2003). *Psychologische Gutachten schreiben und beurteilen*. Vierte, vollständig überarbeitete und erweiterte Auflage. Berlin: Springer.
- Wild, K. P. (1991). *Identifikation hochbegabter Schüler. Lehrer und Schüler als Datenquellen*. Heidelberg: Asanger.