



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



## 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. Übergangsentscheidungen)
  - Systemdiagnostik
    - Zur (vergleichenden) Analyse der Ergebnisse von Bildungsprozessen und –systemen (Output-Diagnostik; z.B. PISA; Lernstandserhebungen)



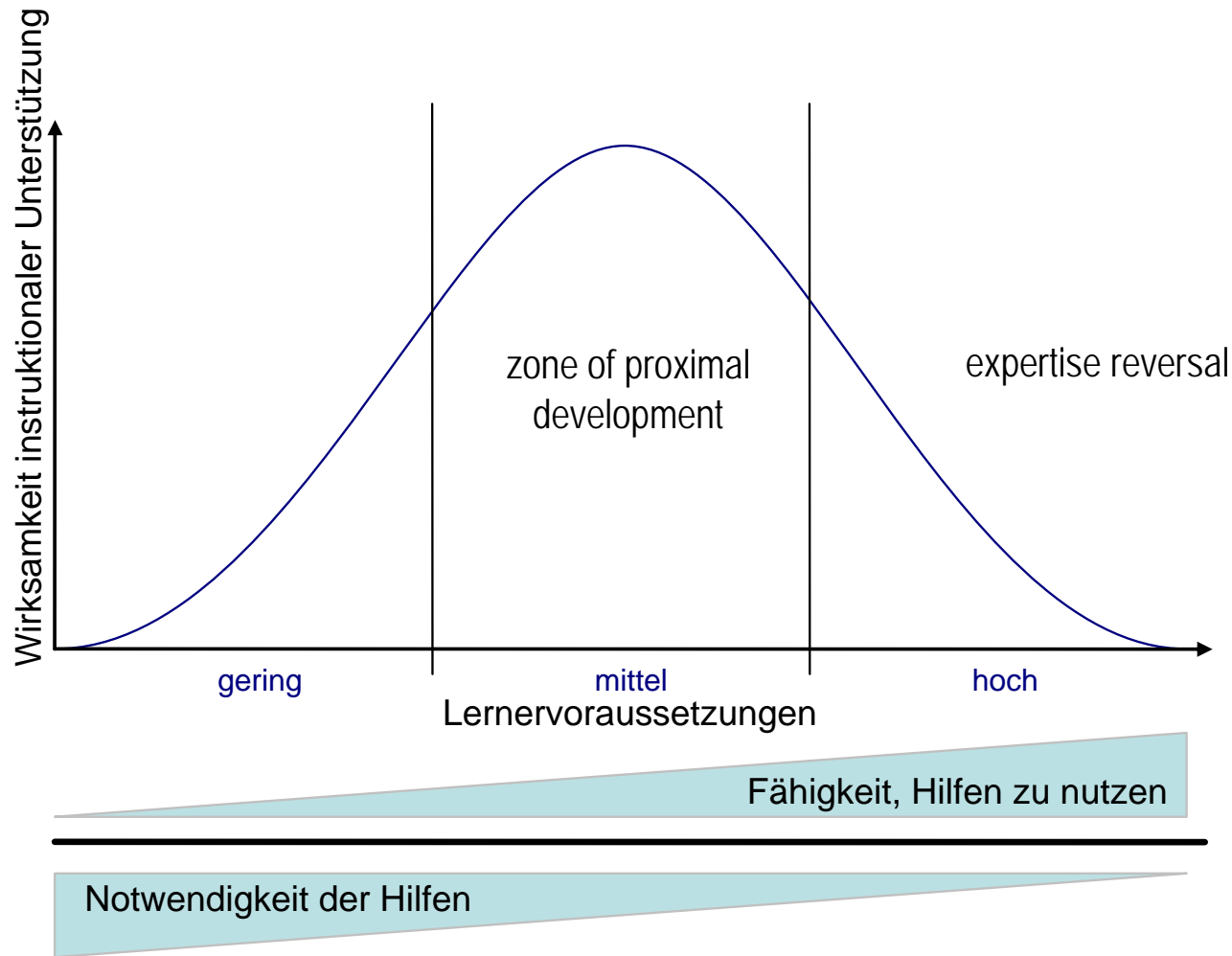
## Funktionen schulischer Diagnostik (2)

- 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



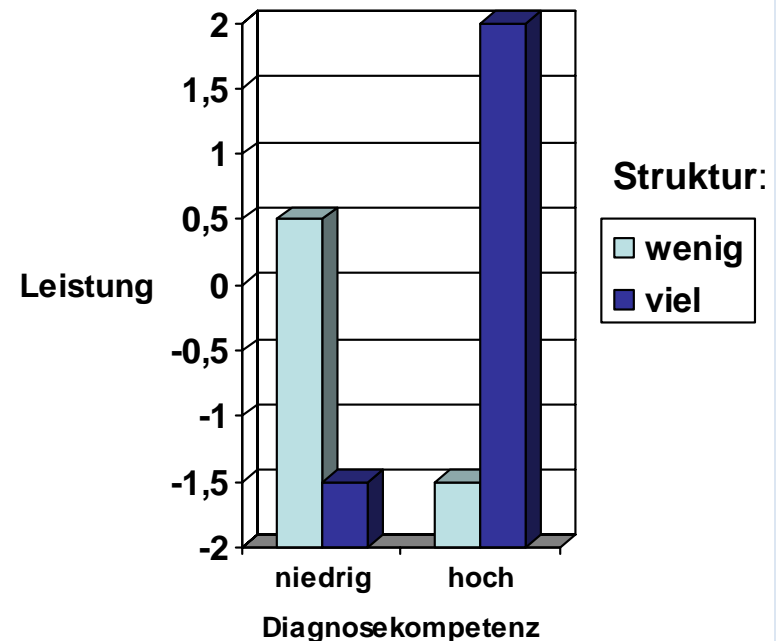
# Beurteilung von Schülerleistungen

- 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 **Differenzierungskomponente** wird eher überschätzt
- Die größte Heterogenität zwischen Lehrern zeigt sich in Bezug auf die **Vergleichskomponente** (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.

## **Diagnostik in der Lehrerbildung**

Vortrag anlässlich des 43. BAK Seminartags in Saarbrücken am 02.10.2009

Univ. Prof. Dr. Roland Brünken  
Lehrstuhl für empirische Bildungsforschung  
FR Erziehungswissenschaft  
Universität des Saarlandes

Meine sehr geehrten Damen und Herren,

ich freue mich sehr über die Einladung zu Ihrem Seminartag und die Möglichkeit, sozusagen vor heimischer Kulisse über die Bedeutung von Diagnostik und diagnostischer Ausbildung in der universitären Lehrerbildung referieren zu können. Aus der Sicht eines Wissenschaftlers und Universitätslehrers – um das vorweg zu schicken – nicht aus der Sicht eines pädagogischen Praktikers möchte ich Ihnen dazu den aktuellen konzeptuellen und empirischen Erkenntnisstand zum Thema Diagnostik in der Lehrerbildung aufzeigen und mögliche Forschungsdesiderate diskutieren.

Ich werde meinen Vortrag dazu grob in folgende Teile gliedern:

Nach einer kurzen Einführung und den unvermeidlichen Blick auf die vieldiskutierten Ergebnisse zur diagnostischen Kompetenz aus der PISA 2000 Studie, möchte ich zunächst versuchen, das Konzept der diagnostischen Kompetenz einzugrenzen und einige methodische Grundüberlegungen anstellen. Danach werde ich Ihnen aktuelle empirische Befunde zum Thema präsentieren. Neben der Frage der Güte diagnostischer Urteile in verschiedenen Urteilsbereichen soll es dabei auch darum gehen zu fragen, in welchem Zusammenhang Diagnostik und Unterrichtserfolg stehen und welche Fördermöglichkeiten wir diesbezüglich sehen. Danach werde ich kurz auf die diagnostischen Anteile in der universitären Lehrerausbildung insgesamt und besonders im Saarland eingehen. Schließen werde ich mit einigen Überlegungen zum Forschungsbedarf und einem kurzen Ausblick.

### Einleitung

Wenn man sich über die Frage unterhält, wie es um die diagnostische Kompetenzen von Lehrern bestellt ist, werden in jeder Diskussion unweigerlich die sattsam bekannten Ergebnisse einer Erhebung aus der PISA Studie 2000 auf den Tisch gelegt, die ich Ihnen auch hier – sozusagen zu Einstimmung – mitgebracht habe. In dieser Erhebung wurde eine kleine Gruppe von Hauptschullehrern gebeten solche Schüler ihrer Klasse zu benennen, von denen Sie vermuten würden, dass es sich um so schwache Leser handelten, dass sie auf Grund ihrer mangelnden Lesekompetenz ernsthafte Schwierigkeiten beim Übergang in den Beruf haben würden. Diese Einschätzung wurde dann verglichen mit den Ergebnissen der Schüler im PISA Lesekompetenztest. Als leseschwach betrachtet wurden dabei solche Schüler, die in PISA nur die unterste Kompetenzstufe erreichten, oder diese sogar verfehlten. Ein Vergleich der PISA Testergebnisse mit den Lehrerurteilen zeigt dabei, dass nur sehr wenige nach PISA leseschwache Schüler von den Lehrern auch als solche benannt wurden.

Die Ergebnisse sind schnell bekannt und viel und kontrovers diskutiert worden. Ich will an dieser Stelle die Diskussion nicht vollständig nachzeichnen, es stellt sich aber aus meiner Sicht die Frage, ob und wie ein Schluss von diesen Ergebnissen auf eine schlechte

diagnostische Kompetenz von Lehrern gerechtfertigt ist und was die Ergebnisse zur Beantwortung der Frage beitragen können, was überhaupt diagnostische Kompetenz meint.

Eine aus meiner Sicht nachvollziehbare und berechtigte Kritik an der vorschnellen Generalisierung der PISA Daten als Ausweis mangelnder diagnostischer Kompetenz von Lehrern findet sich bei Spinath (2004, 2005). Sie kritisiert dabei in erster Linie aus einer methodischen Perspektive heraus, dass die betrachtete Gruppe der Hauptschullehrer nicht repräsentativ sei, dass nur ein Schülermerkmal (nämlich die Lesekompetenz) untersucht wurde und nur ein Akkuratheitsmaß zur Beurteilung der diagnostischen Güte herangezogen wurde (nämlich die Übereinstimmung zwischen Lehrer- und Testurteil). Zudem liege gar kein verbindlicher Vergleichsmaßstab dafür vor, welche Lesekompetenzwerte als „zureichend“ bzw. „unzureichend“ einzuschätzen sind und schließlich blieben widersprechende Forschungsergebnisse unberücksichtigt. Diesen aus meiner Sicht berechtigten Kritikpunkten Spinaths hinzuzufügen ist noch, dass es sich im Falle der PISA-Ergebnisse um eine Messung eines Personenmerkmals und bei der Einschätzung der Lehrer um eine Beurteilung eines Personenmerkmals handelt, was aus psychologischer Perspektive einen fundamentalen Unterschied ausmacht, auf den ich noch näher eingehen werde. Es zeigt sich also, dass die viel diskutierten und Aufsehen erregenden Ergebnisse der PISA Studie nicht dazu geeignet sind, allgemein auf die diagnostischen Kompetenzen von Lehrern zu schließen. Es zeigt sich aber auch an den Daten und insbesondere an den Reaktionen auf diese Ergebnisse, dass es sinnvoll und notwendig ist, wissenschaftlich der Frage nachzugehen, was denn diagnostische Kompetenzen sind und wie es um deren Ausprägung bei Lehrern bestellt ist.

#### Definitionen diagnostischer Kompetenz

Um dieser Frage näher zu kommen, kann man sich zunächst die Standards für die Lehrerbildung der Kultusministerkonferenz (KMK 2004) ansehen. Hierin werden insgesamt 11 Kompetenzen definiert, von denen 2 einen direkten Bezug zu diagnostischen Fragen aufweisen: die Kompetenzen 7 und 8. Kompetenz 7 bezieht sich dabei auf die Diagnostik von Lernvoraussetzungen und Lernprozessen, Kompetenz 8 auf die Diagnostik von Leistungen – also das was wir im schulischen Kontext unter Diagnostik im engeren Sinne – nämlich die Beurteilung in der Regel domänenspezifischer mündlicher und/oder schriftlicher schulischer Leistungen verstehen können.

In wiefern es sich hierbei um eine sinnvolle Trennung zwischen Lernvoraussetzungen, Lernprozessen und Lernergebnissen (Leistungen) handelt, werden wir im Verlauf noch sehen, festzuhalten bleibt zunächst, dass aus politisch-administrativer Sicht es sinnvoll erscheint im Plural von diagnostischen Kompetenzen zu sprechen, und dass diesen im Kontext der professionsbezogenen Kompetenzerwartungen an die Lehrer eine bedeutende Rolle zukommt.

Was sagen nun die Bildungswissenschaften, also in erster Linie die empirische Erziehungswissenschaft und die Pädagogische Psychologie, zur Frage, was man unter diagnostischer Kompetenz bzw. diagnostischen Kompetenzen verstehen kann?

Obschon die Diagnostik insgesamt insbesondere in der Psychologie zu den großen und bedeutenden Fächern gehört, gibt es doch erstaunlicherweise vergleichsweise wenige wissenschaftliche Arbeiten, die sich mit Fragen schulischer Diagnostik befassen. Neben der eben bereits zitierten Arbeit der Kollegin Spinath aus Heidelberg gehört dabei insbesondere die Arbeitsgruppe um Helmke und Schrader aus Landau zu den zumindest auf nationaler Ebene führenden Forschergruppen in diesem Feld.

Nach Schrader (2001) bezeichnet diagnostische Kompetenz allgemein die Fähigkeit, Merkmale von Personen korrekt einzuschätzen. Diagnosen stellen dabei in der Regel explizite, also ausdrückliche, Aussagen über Zustände oder Merkmale von Personen dar, die – und das ist von zentraler Bedeutung – Ergebnisse eines reflektierten und methodisch kontrollierten diagnostischen Prozesses darstellen. Sie resultieren dabei in diagnostischen Urteilen, die wiederum in der Regel der Entscheidungsfindung dienen. Der psychologische Diagnostiker Westhoff redet in diesem Zusammenhang daher auch ganz allgemein von „Entscheidungsorientierter Diagnostik“. (z.B. Westhoff & Kluck, 2003)

Für welche Entscheidungen oder Entscheidungsprozesse werden diagnostische Urteile nun als Kriterien herangezogen? Schrader (2001) unterscheidet hier zwischen einer Mikro- und einer Makroebene. Auf einer Mikroebene dient die Diagnostik einerseits der Unterrichtsplanung, andererseits der individuellen Förderung. Während bei letzterer der einzelne Schüler und seine Leistungsentwicklung im Vordergrund stehen, steht bei Ersterer die mittlere Gruppenleistung im Mittelpunkt. Ihre Erhebung dient der Anpassung von Unterrichtsparametern wie etwa der Stoffmenge, der Instruktionsform oder der Unterrichtsgeschwindigkeit wohingegen die individuelle Leistungsermittlung im Rahmen der Unterrichtsplanung insbesondere für Fragen der Binnendifferenzierung von Bedeutung ist.

Auf einer Makroebene werden vor allem zwei Ziele verfolgt: Unter Selektionsgesichtspunkten dient insbesondere die Leistungsdiagnostik im Sinne einer Prognose zukünftig zu erwartender Leistungsentwicklungen der Vorbereitung und Begründung der individuumsbezogenen Schullaufbahnplanung etwas im Rahmen von Versetzungs- oder Übergangentscheidungen. Unter einer Schulentwicklungsperspektive dient sie aber auch der Systemdiagnostik, also der Beurteilung der Leistungsfähigkeit der Organisation Schule oder gar der Organisationsform Schule, wie wir dies im Rahmen nationaler und internationaler Schulvergleichsstudien, aber auch etwa im Rahmen der Lernstandserhebungen finden.

Es zeigt sich also, dass schulische Diagnostik mit erheblichen Zielerwartungen verbunden ist, und es stellt sich die Frage ob und in wieweit sie diesen, teils konfligierenden Zielen gerecht werden kann.

## Theoretische und Methodische Grundlagen

Grundsätzlich kann man dabei zunächst der Frage nachgehen, wozu überhaupt man Diagnostik im schulischen Kontext betreibt, bzw. betreiben sollte. Auch hierzu gibt Schrader (2001) eine – pädagogische – Begründung: danach besteht die grundsätzliche Idee darin, dass Erziehungsmaßnahmen dann eine optimale Wirkung erzielen, wenn eine hinreichende Passung zwischen den gestellten Anforderungen und den Persönlichkeitsmerkmalen der Betroffenen vorliegt.

Im Rahmen der empirischen Bildungsforschung ist man nun der Frage nachgegangen, ob sich dieser vermutete Zusammenhang auch empirische nachweisen lässt. Die Forschungsrichtung, die sich seit vielen Jahren mit dieser Frage befasst, bezeichnet die Pädagogische Psychologie als Aptitude-Treatment-Interaction-Forschung oder kurz als ATI-Forschung. Diese auf Cronbach (1957) zurückgehende Bezeichnung meint dabei ganz allgemein die Analyse des Zusammenhangs von Persönlichkeitseigenschaften (aptitudes) einerseits und Lehrmethoden (treatments) andererseits. Ausgangspunkt ist dabei der Gedanke, dass nicht jede Lehrmethode in gleicher Weise für alle Personen wirksam sein muss, sondern dass sich deren Wirkung vielmehr zwischen verschiedenen Personen systematisch unterscheiden kann, dass also das, was dem Einen nutzt, für den Anderen nicht notwendigerweise empfehlenswert sein muss.

Wenn dem aber so sein sollte, dann ist es für eine sinnvolle Unterrichtsplanung zunächst notwendig um diese Zusammenhänge zu wissen und entsprechende Lernvoraussetzungen mit in den Blick zu nehmen. Dies setzt wiederum eine funktionierende Diagnostik voraus, und zwar sowohl – Sie erinnern sich an die Unterscheidung der Kompetenzbereiche 7 und 8 in der KMK Standards - hinsichtlich der Lernvoraussetzungen, als auch hinsichtlich der Lernergebnisse.

Ein besonders plastisches Beispiel für den Zusammenhang von Lernvoraussetzungen und Lehrmethode liefert dabei die Forschung zum sogenannten „Expertise-Reversal-Effekt“ (Kalyuga et al, 2003). Kalyuga ist hierbei in verschiedenen Studien dem Zusammenhang individuellen, domänenspezifischen Vorwissens mit Methoden des angeleiteten Lernens, z.B. dem Lernen aus Lösungsbeispielen, nachgegangen. Hierbei zeigte sich ein eindeutiger, aber eben nicht linearer Zusammenhang zwischen Expertise und Anleitung. Demnach hängt die Wirkung angeleiteten Lernens vom Grad der Expertise in der Form ab, dass mit zunehmender Expertise die Lerner zunächst von der Unterstützung durch den Lehrer bzw. die Lehrmethode immer stärker profitieren. Während Anfänger zunächst noch wenig Leistungssteigerung zeigen, wächst die Effizienz mit zunehmenden Vorkenntnissen. Dieser positive Effekt nimmt jedoch mit wachsender Expertise nicht weiter zu, sondern im Gegenteil bei weiter wachsender Expertise profitieren die Lerner von der Unterstützung nicht nur nicht mehr, sie werden sogar vergleichsweise schlechter, der anfangs positive Effekt dreht sich also in einen negativen Effekt um, daher die Bezeichnung Expertise-Reversal-Effekt. Dieser Effekt konnte in einer Vielzahl von Studien, auch von unserer eigenen Arbeitsgruppe (Brünken & Leutner, 2005; Seufert, Jänen & Brünken, 2007) nachgewiesen werden und gilt als empirisch gut gesichert. Wir gehen davon aus, dass hier zwei unterschiedliche Aspekte sich gegenseitig beeinflussen: einerseits die mit zunehmender Expertise sinkende Notwendigkeit, Hilfen zu nutzen, andererseits die damit steigenden Fähigkeiten, angebotene Hilfen auch effizient nutzen zu können. Im Sinne Vygotskys kann man dies auch als Ausweis einer Zone proximaler Entwicklung interpretieren.

Kehren wir zurück zu unserer Ausgangsfrage der schulischen Diagnostik. Das Beispiel soll zeigen, in welchem Zusammenhang Diagnostik und erfolgreiche Unterrichtsplanung stehen, setzt doch hier ein wirkungsoptimierter Einsatz instruktorischer Unterstützung voraus, dass ich den Expertisestand meiner Schüler korrekt diagnostizieren kann. Hierbei handelt es sich um eine typische diagnostische Anforderung an Lehrende, nämlich um eine (informelle) Leistungsdiagnostik. Neben dieser sicher besonders verbreiteten Form schulischer Diagnostik gibt es noch eine Reihe weiterer typischer diagnostischer Anforderungen im schulischen Kontext. Dazu gehören die Verhaltensdiagnostik, der in den letzten Jahren zunehmende Bedeutung etwa im Bereich der Grundschule zugekommen ist, der Motivations- und Volitionsdiagnostik, aber auch der Verdachtsdiagnostik von Leistungsauffälligkeiten, sowohl im unteren, als auch im oberen Leistungsbereich, sowie der Diagnostik von Verhaltensauffälligkeiten. Typisch für all diese Formen von Diagnostik ist, dass dem Lehrenden dazu in der Regel kein standardisiertes diagnostisches Instrumentarium, kein „Fieberthermometer“ und kein allgemein verwendbarer Test zur Verfügung steht. D.h. neben der eigentlichen diagnostischen Anforderung steht häufig noch die Frage nach den zu verwendenden diagnostischen Methoden.

Dem entgegen steht die so genannte standardisierte pädagogisch-psychologische Diagnostik, die durchaus über solche Methoden verfügt – etwa zur Intelligenzdiagnostik, aber auch zur Diagnostik von Teilleistungsstörungen oder auch zur klinischen Verhaltensdiagnostik. Deren Anwendung und Auswertung setzt aber in aller Regel eine spezielle Ausbildung voraus und ist nicht Aufgabe von Lehrern.



## Gütekriterien diagnostischer Urteile

Auch wenn es im Bereich schulischer Diagnostik solche Instrumente (noch) nicht gibt, so gibt es doch auch hier unbestrittene Gütekriterien, nach denen man die Qualität diagnostischer Prozesse bewerten kann. Hierbei handelt es sich im Übrigen um die gleichen Kriterien, die auch bei standardisierten diagnostischen Verfahren zur Anwendung kommen. Die drei wichtigsten dieser Kriterien sind: die Validität (oder Gültigkeit), die Reliabilität (oder Zuverlässigkeit) und die Objektivität. Helmke et al (2004) führen dem noch ein viertes, spezifisches Kriterium hinzu, die Veridikalität: gemeint ist hiermit die Übereinstimmung des Lehrerurteils mit einem objektiv – beispielsweise durch einen Schulleistungstest gemessenen Merkmal. In wieweit es sich hierbei um ein eigenständiges Gütekriterium oder um eine Variante des Kriteriums Validität handelt, kann man sicher trefflich diskutieren. Wichtiger als diese eher akademische Kontroverse ist in unserem Zusammenhang die Frage, welche Faktoren denn die Güte eines diagnostischen Urteils überhaupt beeinflussen.

Schrader (1989) identifiziert hierbei drei Aspekte der Urteilsgenauigkeit:

- (1) das diagnostische Wissen des Lehrenden hinsichtlich der relevanten zu beurteilenden Merkmale;
- (2) die anzuwendenden Urteilsmaßstäbe, die je nach diagnostischem Ziel verschieden sein können und müssen. Wir unterscheiden hierbei individuelle, soziale und kriteriale (oder curriculare) Bezugsnormen, die man zur Beurteilung einer individuellen Merkmalsausprägung heranziehen kann. Während man beispielsweise zur Unterrichtsplanung in der Regel die soziale Bezugsnorm verwendet, dient die individuelle Bezugsnorm der Förderdiagnostik und die kriteriale Bezugsnorm der schulischen Leistungsdiagnostik.
- (3) Schließlich werden diagnostische Urteile von Urteilstendenzen beeinflusst. Hierbei handelt es sich um typische, schemagetriebene Urteilsfehler, die in der Regel dadurch entstehen, dass bestimmte Randinformationen den Urteilsprozess mit beeinflussen obwohl sie in keinem oder zumindest keinem direkten Zusammenhang zum zu beurteilten Merkmal stehen. So tendieren Lehrende beispielsweise dazu in den Fällen, in denen Sie Schüler in mehreren Fächern unterrichten, die Leistungen der Schüler fächerübergreifend zu vergleichen, was unter Umständen zu Milde- oder Strengefehlern führen kann.

Wie nun kann man beurteilen, wie gut ein diagnostisches Urteil ist? In der Regel werden hierzu so genannte Akkuratheitsmaße verwendet, bei denen die Übereinstimmung des Lehrerurteils mit einem Außenkriterium oder die Vorhersagegenauigkeit bestimmt werden. Helmke und Schrader (Schrader & Helmke, 1987; Helmke et al, 2004) haben hierzu eine sinnvolle Differenzierung vorgeschlagen. Demnach kann man bei diagnostischen Urteilen von Lehrern drei Komponenten unterscheiden, eine Niveau- eine Differenzierungs- und eine Vergleichskomponente:

- hinsichtlich der Niveauebene geht es um die Einschätzung der mittleren Leistungsfähigkeit einer Gruppe;
- in Bezug auf die Differenzierungskomponente um die Leistungshomogenität der Gruppe.
- die Vergleichskomponente, die Helmke und Schrader als das diagnostische Urteil im engeren Sinne ansehen, bezieht sich nun auf den Vergleich der Schüler untereinander.

## Empirische Forschungsergebnisse

Was sagt nun die empirische Forschung zur Urteilsgenauigkeit von Lehrerurteilen? Insgesamt ist die empirische Basis hierzu noch vergleichsweise dünn, es zeigen sich aber einige stabile Befundmuster. In der Regel verwenden die Studien dabei ein oder mehrere Akkuratheitsmaße, wobei normalerweise Schülern Aufgaben vorgelegt werden und die Lehrer einschätzen sollen, ob und in wieweit die Schüler dazu in der Lage sind, die Aufgaben zu lösen. Vergleichen wird dann das Lehrerurteil mit der tatsächlichen Aufgabenlösung. Ähnlich werden Leistungsvoraussetzungen durch den Vergleich des Lehrerurteils mit standardisierten Tests untersucht. Neben der Schülerleistung werden dabei in der Regel solche Leistungsvoraussetzungen wie z.B. Intelligenz, Motivation, Fähigkeitsselbstkonzept und Leistungsängstlichkeit betrachtet, die in einem engen Zusammenhang mit schulischer Leistung stehen.

Welche Ergebnisse zeigen solche Studien?

- Lehrer neigen zu einer (leichten) Niveauüberschätzung (Schrader & Helmke, 1987), die von normativen Aspekten (was müßten Schüler können?) getragen wird und eher kompetenz- als Leistungsorientiert ist (was könnten die Schüler *eigentlich* leisten?).
- Lehrer neigen dazu, Unterschiede zu akzentuieren, d.h. die Differenzierungskomponente wird eher überschätzt
- Die größte Heterogenität zwischen Lehrern zeigt sich in Bezug auf die Vergleichskomponente (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)
- Die durchschnittliche mittlere Übereinstimmung zwischen Leistungstest und Lehrerurteil liegt dabei bei  $r = .66$ , einem mittleren Zusammenhang (Metaanalyse von Hoge & Coladarci, 1989).

Interessant in diesem Zusammenhang ist eine schon ältere Studie (Weinert & Helmke, 1987), die einen Zusammenhang von diagnostischer Kompetenz einerseits und Unterrichtserfolg andererseits nachweisen konnte:

demnach erleichtert eine hohe diagnostische Kompetenz es den Lehrern, lernförderliche unterstützende Maßnahmen zielgerichtet einzusetzen, wohingegen sich bei Lehrern mit geringer diagnostischer Kompetenz Unterstützungsmaßnahmen als ineffizient (Fehlallokation von Förderung) erweisen. Diese Ergebnisse passen sehr gut zu den eingangs gezeigten Ergebnissen zum Expertise-Reversal-Effekt, der einen ähnlichen Zusammenhang aus Lernerperspektive zeigt.

Hinsichtlich der Beurteilung von Lernvoraussetzungen ist insbesondere die Intelligenzdiagnostik in den Blick genommen worden. Hierbei hat vor allem Wild (1992) im Rahmen des Marburger Hochbegabtenprojektes interessante Ergebnisse gefunden. Dabei zeigte sich insgesamt eine gute mittlere Übereinstimmung zwischen Lehrereinschätzung und allgemeiner Intelligenz ( $r = .67$ ), wobei die Übereinstimmung bei der Einschätzung spezifischer Faktoren (verbal, numerisch, räumlich) geringer war. Aber es zeigte sich auch, dass die Lehrereinschätzung hoch leistungskorreliert war, d.h. die Intelligenzhöhe wird aus hoher Leistung geschlossen. Daher haben Lehrende Probleme bei der Identifikation von so genannten Underachievern, also solchen Lernern die bei hohen psychometrischen IQ – Werten nur durchschnittliche bis niedrige schulische Leistungen erbringen.

In Bezug auf die Einschätzung zu nicht kognitiven Lernvoraussetzungen gibt es nur sehr wenige Studien. Als Akkuratheitsmaß wird hierbei die Korrelation zwischen dem Lehrerurteil und der Schülerelbsteinschätzung berechnet. Einige ausgewählte Ergebnisse hierzu sind, das Lehrer das Fähigkeitsselbstkonzept der Lernenden relativ gut einschätzen können (mittlere Zusammenhänge  $.30 < r < .60$ ; Marsch, 1990), wohingegen es nur wenig Übereinstimmung in Bezug auf die Lernmotivation gibt (geringe positive Zusammenhänge;  $r < .30$ ; Helmke & Fend, 1981). Auch in Bezug auf die Leistungsängstlichkeit gibt es nur schwache positive Zusammenhänge zwischen Lehrer- und Schülerelbsteinschätzung ( $r < .30$ ; Böhnke et al, 1986). Insgesamt zeigen sich hier wiederum Hinweise auf den Einfluss von Leistungen auf die Einschätzung (z.B. werden gute Schüler als motivierter angesehen als weniger gute).

In einer neueren Studie ist Spinath (Spinath, 2005) der Frage nachgegangen, ob es sich bei der diagnostischen Kompetenz von Lehrenden überhaupt um ein eindimensionales Konstrukt handelt oder ob hier nicht verschiedene Fähigkeiten „in einen Topf“ geworfen werden. Die Ergebnisse zeigen dabei in Bezug auf die drei von Helmke und Schrader vorgeschlagenen Facetten nur geringe Zusammenhänge, ebenso gibt es deutliche Unterschiede zwischen Leistungsdiagnostik und der Diagnostik nicht-kognitiver Aspekte. Insgesamt folgert sie daraus, dass es nicht „die diagnostische Kompetenz“ gibt.

Zur Frage, ob und wie man die diagnostische Kompetenz von Lehrenden Fördern kann schließlich gibt es so gut wie keine empirischen Daten. Zwar liegen einige pragmatische Empfehlungen etwa zur Verwendung von Daten aus schulübergreifenden Lernstandserhebungen (Helmke et al, 2004), zur Schaffung von Übungsmöglichkeiten (Wahl, Weinert & Huber, 1997) und zur Auswahl richtiger Beobachtungsmerkmale für diagnostische urteile (Kleber, 1992), es fehlen aber vollständig längsschnittliche Studien zur Kompetenzentwicklung.

#### Zusammenfassung empirischer Befunde

Zusammenfassend lässt sich der Stand der empirischen Forschung also so beschreiben:

- Die Leistungsdiagnostik funktioniert relativ gut bei einer Tendenz zur Verwendung sozialer Bezugsnormen, einer Tendenz zur generellen Überschätzung der Schülerleistungen (Niveauelemente) und weitgehend fehlenden externen Bezugskriterien
- Auch die Intelligenzdiagnostik von Lehrenden ist relativ zuverlässig, aber sie ist hoch leistungskorreliert was zu Probleme bei der Identifikation von Underachievern führt.
- Hinsichtlich nicht kognitiver Schülermerkmale ist die Beurteilungsgüte umso besser, je enger das Merkmal mit Leistung korreliert ist; dabei wird insbesondere die Leistungsängstlichkeit eher unterschätzt.
- 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 Lehrerbildung

Bei der Neustrukturierung des bildungswissenschaftlichen Anteils an der Lehrerbildung haben wir daher versucht, dem Bereich der Diagnostik breiteren Raum als bisher einzuräumen, wobei wir insbesondere – im Rahmen der universitären Ausbildung, auf die Aspekte diagnostisches Wissen und diagnostische Methoden fokussiert haben. Das Saarbrücker Modell der Bildungswissenschaften in der Lehrerbildung (UdS SPO 2007) stellt einen

modularisierten, an den KMK Richtlinien und der Empfehlungen von DGfE und DGPs orientierten Studiengang mit 48 CP (>25% der Lehramtsausbildung) dar.

Wir haben dabei 4 Kompetenzbereiche in den Mittelpunkt gestellt:

- K1: Lehren und Lernen
- K2: Persönlichkeitsentwicklung und Erziehung
- K3: Diagnostik, Beratung, Intervention
- K4: Schulentwicklung, Qualitätssicherung & Systemmonitoring

Im Mittelpunkt der diagnostischen Ausbildung steht dabei natürlich der Kompetenzbereich 3, zu dessen Inhalte gehören:

- eine einführende Vorlesung in die Diagnostik
- Seminare und Übungen zu schulische Leistungsdiagnostik
- Seminare und Übungen zur Förderdiagnostik
- sowie zur standardisierten Diagnostik (psychometrische Testverfahren)

Die Diagnostik ist aber auch in den anderen Kompetenzbereichen vertreten, z.B.

- mit der Vorlesung zu bildungswissenschaftlichen Grundlagen (K1)
- mit einer Einführung in die empirischen Forschungsmethoden (K1)
- mit Motivation & Motivationsdiagnostik (K1)
- mit der Diagnostik von Schulschwierigkeiten und Disziplinproblemen (K2)
- sowie der (international) Vergleichende Bildungssystemdiagnostik (K4)

Insgesamt macht die diagnostische Ausbildung also einen Kernbereich der universitären Lehrerausbildung aus.

Wir sehen also in der modernen Lehrerausbildung bereits einige Fortschritte in Bezug auf die Entwicklung diagnostischer Kompetenzen bei den Lehramtsstudierenden. Aber in vielen Bereichen besteht doch noch ein erheblicher Forschungsbedarf. Zur Kompetenzdiagnostik fehlen derzeit noch brauchbare, schulfachbezogene Kompetenzmodelle und darauf basierende Testverfahren. Hier ist jedoch mit dem neuen DFG Schwerpunktprogramm "Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen" (Klieme & Leutner, 2008) Abhilfe in Sicht. Ebenso sind die individuelle Voraussetzung erfolgreicher Lehramtsstudierender noch unklar. Hier gibt es derzeit ein recht umfangreiches Programm des BMFT zur Förderung von Forschung im Bereich der Professionalisierung der Lehrerbildung. An diesem Programm sind wir hier in Saarbücken mit dem Projekt SIOS-L (Brünken, Spinath & Bedersdorfer, 2009) beteiligt, in dem wir eine auch längsschnittliche Studie zur Entwicklung pädagogischer Kompetenzen durchführen werden (übrigens unter Beteiligung aller an der saarländischen Lehrerausbildung beteiligten Institutionen). Schließlich fehlen ebenfalls noch weitgehend Evaluationsstudien zur Effizienz von Lehreraus- und weiterbildung.

Ausblick

Zusammenfassen kann man also festhalten, dass eine generelle Kritik an der diagnostischen Kompetenz von Lehrern nicht gerechtfertigt ist. Insbesondere im Leistungsbereich sind Lehrer zuverlässige, wenn auch in der Regel „milde“ Diagnostiker. Diagnostische Urteile von Lehrern haben dabei in allen Bereichen einen „Leistungs-Bias“; d.h. im Mittelpunkt aller diagnostischen Einschätzungen von Lehrern steht die schulische Leistung der Schüler. Von

ihr aus wird auf Lernvoraussetzungen ebenso wie auf die Effizienz von Lernprozessen geschlossen.

Die universitäre diagnostische Ausbildung dient in erster Linie der Vermittlung von

- diagnostischem Wissen
- Diagnostischen Methoden und Techniken
- Grundlagen der Urteilsbildung und ihrer Schwierigkeiten
- Und Grundlagen der standardisierten Diagnostik

Das Konzept der diagnostischen Kompetenz schließlich als eindimensionales Konstrukt ist umstritten; insgesamt ist hier eine differenziertere Perspektive einzunehmen.

Ich danke Ihnen für ihre Aufmerksamkeit!

### **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.
- 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.