

Schrader, Friedrich-Wilhelm

## Diagnostische Kompetenz von Lehrpersonen

*Beiträge zur Lehrerinnen- und Lehrerbildung 31 (2013) 2, S. 154-165*



Empfohlene Zitierung/ Suggested Citation:

Schrader, Friedrich-Wilhelm: Diagnostische Kompetenz von Lehrpersonen - In: Beiträge zur Lehrerinnen- und Lehrerbildung 31 (2013) 2, S. 154-165 - URN: urn:nbn:de:0111-pedocs-138435

in Kooperation mit / in cooperation with:

Zeitschrift zu Theorie und Praxis der Aus- und  
Weiterbildung von Lehrerinnen und Lehrern

BEITRÄGE ZUR LEHRERINNEN-  
UND LEHRERBILDUNG

Organ der Schweizerischen Gesellschaft für  
Lehrerinnen- und Lehrerbildung (SGL)

ISSN 2296-9632

<http://www.bzl-online.ch>

### Nutzungsbedingungen

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

### Terms of use

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

By using this particular document, you accept the above-stated conditions of use.

### Kontakt / Contact:

peDOCS  
Deutsches Institut für Internationale Pädagogische Forschung (DIPF)  
Informationszentrum (IZ) Bildung  
E-Mail: [pedocs@dipf.de](mailto:pedocs@dipf.de)  
Internet: [www.pedocs.de](http://www.pedocs.de)

Mitglied der

  
Leibniz-Gemeinschaft

## Diagnostische Kompetenz von Lehrpersonen\*

Friedrich-Wilhelm Schrader

**Zusammenfassung** Der Beitrag gibt einen Überblick über Diagnoseleistungen und diagnostische Kompetenzen in der Lehrtätigkeit. Zu Beginn werden Grundlagen und basale Aufgabenstellungen der pädagogischen Diagnostik behandelt. Im Vordergrund steht dabei die Bedeutung der pädagogischen Diagnostik für den Lehr-Lern-Prozess. Einen weiteren Schwerpunkt bilden der Ansatz der Urteilsgenauigkeit sowie prozessorientierte Ansätze, zu denen jeweils ein knapper Überblick über ausgewählte Forschungsergebnisse gegeben wird. Am Ende wird auf den Stellenwert diagnostischer Kompetenz, Forschungsdefizite sowie Konsequenzen für die Verbesserung von Diagnoseleistungen eingegangen.

**Schlagwörter** diagnostische Kompetenz – Diagnosegenauigkeit – formelle und informelle Diagnosen – Urteilsprozesse – Unterricht

### Teacher Diagnosis and Diagnostic Competence

**Abstract** The following article provides an overview of teacher diagnosis and diagnostic competence. Starting with a presentation of basic approaches and diagnostic tasks, the main focus lies on the role of educational diagnosis in instruction and with respect to student learning. Special emphasis is put on two major approaches: teachers' judgement accuracy and judgement processes. A brief review of selected results of research in these domains is given. Finally, the importance of diagnostic competence, limitations of research, and consequences regarding the improvement of teacher diagnosis are discussed.

**Keywords** diagnostic competence – judgement accuracy – formal and informal diagnosis – judgement processes – instruction

## 1 Einleitung

Diagnostische Kompetenz von Lehrpersonen bezieht sich auf die Fähigkeit, die im Lehrberuf anfallenden diagnostischen Aufgabenstellungen erfolgreich zu bewältigen, und auf die Qualität der dabei erbrachten Diagnoseleistungen (Schrader, 2011). Ziel ist es, Informationen über Lernergebnisse, Lernvoraussetzungen und Lernvorgänge von Schülerinnen und Schülern zu gewinnen, die für verschiedene pädagogische Entscheidungen (Notengebung, Versetzung, Übergangsempfehlungen, Unterrichtsplanung und -gestaltung, Schul- und Unterrichtsentwicklung) genutzt werden können. Diagno-

---

\* Ich danke Andreas Helmke und zwei unbekanntem Gutachtern für hilfreiche Kommentare zu einer ersten Fassung dieses Beitrags.

sen betreffen den Ist-Zustand und stehen in Bezug zu bestimmten Zielvorstellungen (Soll-Zustände), die sich aus Lehrzielen, curricularen Vorgaben und Bildungsstandards ableiten. Schwerpunkt der Lehrpersonendiagnostik sind Leistungen der Schülerinnen und Schüler, die in der schulischen Praxis das zentrale Zielkriterium des Unterrichts darstellen und gleichzeitig grundlegende Lernvoraussetzungen sind. Während früher Fragen der Beurteilung und Bewertung von Lernergebnissen im Vordergrund standen, richtet sich der Blick gegenwärtig stärker auf die Nutzung der Diagnostik für die Unterrichtsgestaltung, die Steuerung des Lehr-Lern-Prozesses und die Unterrichtsentwicklung. Diagnostische Kompetenz wird häufig mit der Fähigkeit gleichgesetzt, genaue diagnostische Urteile abzugeben (Schrader, 2011). Daneben rücken aber Fragen, die den Urteilsprozess und die Urteilsbildung betreffen, zunehmend in den Blickpunkt wissenschaftlichen Interesses (Artelt & Gräsel, 2009).

## **2 Grundlagen des pädagogischen Diagnostizierens von Lehrpersonen**

Während in der pädagogischen Diagnostik wissenschaftlich fundierte Methoden des Diagnostizierens (formelle Diagnostik) im Vordergrund stehen, geht es bei der diagnostischen Kompetenz von Lehrpersonen stärker um die im Schul- und Unterrichtsalltag vorherrschenden Urteile. Diese informellen (bzw. semiformalen, vgl. Hascher, 2008) Diagnoseleistungen sind für das unterrichtliche Handeln zwar unverzichtbar, wegen ihrer Fehleranfälligkeit jedoch problematisch, sodass oft zu fragen ist, inwieweit sie durch formelle Diagnosen ergänzt oder korrigiert werden können.

Die *abschliessende Bewertung von Lernprozessen* ist Grundlage für Versetzungs- und Übertrittsentscheidungen und für die Vergabe von Berechtigungen (Abschlusszeugnisse), die oft für Selektionszwecke (Ausbildungs-, Studien- oder Arbeitsplätze) herangezogen werden und langfristige und gravierende Folgen für die Einzelne und den Einzelnen wie auch für die Gesellschaft haben (Ingenkamp & Lissmann, 2008; van Ophuysen, 2010). Deshalb sind Forderungen nach einer pädagogischen Diagnostik, die hohen Gütemassstäben (Objektivität, Reliabilität, Validität) gerecht wird, und nach einer fundierten methodischen Absicherung der Urteile gerade hier von grosser Bedeutung. Ein weiterer Schwerpunkt der Diagnostik sind *unterrichtsbezogene Diagnoseleistungen*, mit denen der Lehr-Lern-Prozess überwacht und optimiert wird (Ingenkamp & Lissmann, 2008; Nitko, 2004). Grundlegend für den Lern- und Unterrichtserfolg ist eine hinreichende Anpassung des Unterrichts an die Lernvoraussetzungen der Schülerinnen und Schüler (Adaptivität), die sowohl durch längerfristige, auf grössere Unterrichtseinheiten bezogene Massnahmen (*Makroadaptationen*) als auch durch kurzfristige Abstimmungsmassnahmen während des Unterrichts (*Mikroadaptationen*) erfolgen kann (vgl. z.B. Schrader & Helmke, 2001). Während Mikroadaptationen meist nur auf der Grundlage von diagnostischen Informationen (Beobachtungen, subjektive Einschätzungen), die während des Unterrichtens anfallen, erfolgen können, besteht bei

Makroadaptationen oftmals die Möglichkeit, gezieltere diagnostische Abklärungen (formelle Diagnostik) vorzunehmen.

### **Abklärung der Lernvoraussetzungen**

Eine systematische Abklärung der Lernausgangslage ist vor allem bei der Übernahme einer neuen Klasse oder nach dem Wechsel in einen weiterführenden Bildungsgang sinnvoll und kann wichtige Hinweise für die Planung und Gestaltung des Unterrichts liefern (Hesse & Latzko, 2009). Schwerpunkt sind bereichsspezifische Vorkenntnisse, die sich als die besten Prädiktoren des späteren Lernerfolgs erwiesen haben (Helmke & Weinert, 1997). Vor allem in hierarchisch aufgebauten Fächern wie Mathematik handelt es sich um kumulative Lernergebnisse, bei denen einmal entstandene Lernlücken das nachfolgende Lernen (und die Lernmotivation) grundlegend beeinträchtigen können. Diagnostische Hilfsmittel sind Strukturanalysen des Vorwissens, die auf einer detaillierten, die herkömmlichen Leistungsüberprüfungen (Klassenarbeiten) ergänzenden Dokumentation von Lernverläufen basieren (Hesse & Latzko, 2009).

### **Überwachung des Lernfortschritts**

Traditionelle sozialnorm- oder bezugsgruppenorientierte Schulleistungstests bieten die Möglichkeit, den Lernstand im Vergleich zur gesamten Altersgruppe festzustellen. Da sie meistens an Groblernzielen orientiert und nicht auf den Unterricht in einer bestimmten Klasse bezogen sind, geben sie oft nur wenige Hinweise für die Gestaltung des weiteren Lernprozesses. Kriteriumsorientierte Tests dagegen sind an Feinlernzielen orientiert und erlauben es, das Erreichen der Lehrziele an einer sachlichen Bezugsnorm zu messen (Ingenkamp & Lissmann, 2008). Für die eigentliche Lernsteuerung, bei der eine hohe Messgenauigkeit nicht im Vordergrund steht, kommen darüber hinaus von den Lehrpersonen entwickelte informelle Verfahren, offene Angaben sowie Interview- und Beobachtungsformate infrage (Shepard, 2001).

### **Abklärung von Lernschwierigkeiten**

In einigen Bereichen – beispielhaft ist hier insbesondere die Lesekompetenz zu nennen – bieten grundlagenwissenschaftlich fundierte Testverfahren die Möglichkeit, Teilkomponenten von fachlichen Fähigkeiten zu erfassen (Hasselhorn, Schneider & Marx, 2000). Für eine breitere Ursachenanalyse, die auch den nicht kognitiven Bereich (insbesondere Lernmotivation und Lernverhalten) einbezieht, sind mittlerweile wissenschaftliche Ansätze und Messverfahren verfügbar (vgl. z.B. Hesse & Latzko, 2009), die sehr aussichtsreich erscheinen, auf die Lehrpersonen in ihrer Ausbildung aber bislang noch kaum vorbereitet werden.

### **Unterrichtsentwicklung und Professionalisierung**

Für die Bestandsaufnahme bieten schul- und klassenübergreifende Kompetenzmessungen (Vergleichsarbeiten, andere Lernstandserhebungen) besondere Vorteile, da sie anders als viele herkömmliche Schulleistungstests eng an Kernkompetenzen bzw. Bildungsstandards orientiert sind und oft innovative Aufgabenstellungen umfassen, die

für den weiteren Unterricht wichtige Impulse liefern können (Altrichter, 2010; Helmke & Hosenfeld, 2004). Das Verständnis und die Fähigkeit zum Umgang mit solchen Informationen bilden einen zunehmend wichtigen Aspekt der diagnostischen Kompetenz. Lohnenswert erscheint es, über reine Outputdaten hinaus auch Kontext- und Prozessdaten einzubeziehen (etwa durch die Nutzung von Schüler- und kollegialem Feedback; vgl. Helmke et al., 2012).

### 3 Empirische Forschungen zur diagnostischen Kompetenz

#### 3.1 Diagnostische Kompetenz als Diagnosegenauigkeit

Die seit Langem bekannten Mängel der Notengebung (unzureichende Objektivität, Reliabilität, Validität; vgl. z.B. Ingenkamp & Lissmann, 2008) werden oft als Beleg für die mangelnde diagnostische Kompetenz von Lehrpersonen angeführt. Der Ansatz der Urteilsgenauigkeit beschäftigt sich systematisch mit der Frage, wie gut Lehrpersonenurteile mit den bei den Schülerinnen und Schülern gemessenen «tatsächlichen» Merkmalsausprägungen übereinstimmen. Sofern Lehrpersonenurteile und Merkmale der Schülerinnen und Schüler im gleichen Antwortformat erfasst werden (z.B. Einschätzung der von jeder Schülerin und von jedem Schüler in einem Test erreichten Punktzahl), sind alle Abweichungen zwischen beiden als Fehler zu betrachten. Bei einfachen Abweichungsmassen werden aber verschiedene Komponenten der Urteilsgenauigkeit vermischt (Cronbach, 1955), sodass üblicherweise Komponentenanalysen durchgeführt werden. In der deutschsprachigen Forschung geschieht dies allerdings fast ausschliesslich in einer vereinfachten Form, bei der eine Urteilerin oder ein Urteiler (z.B. eine Lehrperson) mehrere Personen (z.B. die Schülerinnen und Schüler der eigenen Klasse) hinsichtlich *eines* einzigen Merkmals (statt mehrerer wie bei Cronbach, 1955) beurteilt (Karst, 2012). Dabei können eine (jeweils klassenweise berechnete) *Niveauelemente* (Über- oder Unterschätzung des mittleren Ausprägungsniveaus), eine *Differenzierungs- oder Streuungskomponente* (Über- oder Unterschätzung der Merkmalsstreuung) und eine *Rangordnungskomponente* (Korrelation zwischen Urteilen und tatsächlichen Ausprägungen) unterschieden werden (Schrader & Helmke, 1987). Die Rangordnungskomponente wird häufig als das eigentliche Mass für die diagnostische Kompetenz (diagnostische Sensitivität) angesehen, da sie von allgemeinen Urteilstendenzen (wie Milde- oder Strengetendenz) frei ist und sich auch berechnen lässt, wenn Urteile und tatsächliche Merkmalsausprägungen nicht im gleichen Antwortformat erfasst werden. Gelegentlich werden die Leistungen der Schülerinnen und Schüler auch auf einzelne Aufgaben bezogen beurteilt (gelöst/nicht gelöst) und Treffermasse berechnet (Coladarci, 1986; Karing, Matthäi & Artelt, 2011). Am häufigsten untersucht ist die Beurteilung von Leistungen der Schülerinnen und Schüler, die nicht nur eine zentrale Rolle für den Unterricht spielen, sondern sich mithilfe von Tests auch genauer und weniger verzerrungsanfällig messen lassen als nicht kognitive Merkmale der Schülerinnen und Schüler (z.B. Interesse: Karing, 2009; Fähigkeitsselbstkonzept: Praetorius, Greb, Dickhäuser & Lipowsky, 2011).

### 3.2 Deskriptive Ergebnisse zur Diagnosegenauigkeit

Aus Platzgründen kann hier nur ein knapper Überblick über einige zentrale Befunde zur personenbezogenen diagnostischen Kompetenz gegeben werden (zur Einschätzung von Aufgabenschwierigkeiten vgl. McElvany et al., 2009). So zeigen verschiedene Studien, dass Lehrpersonen im Durchschnitt gesehen die Rangordnung der Leistungen recht genau einschätzen (Hoge & Coladarsi, 1989). In einer auf der Basis von 75 Studien durchgeführten mehrebenenanalytisch ausgewerteten Meta-Analyse von Südkamp, Kaiser und Möller (2012) wird eine mittlere Korrelation von .63 bei einer Spannweite von -.03 bis .84 berichtet. Signifikante Moderatoren sind lediglich Testkenntnis und Übereinstimmung im Spezifitätsgrad von Urteilen und Leistungen; andere Faktoren, darunter das Fachgebiet (Sprachleistungen versus Mathematik), spielen hingegen keine Rolle. Zwischen einzelnen Lehrpersonen bzw. Klassen gibt es deutliche Unterschiede in der Höhe der Korrelation, z.B. von .02 bis .88 bei Helmke und Schrader (1987) oder von .11 bis .86 bei Hosenfeld, Helmke und Schrader (2002), und damit in der Genauigkeit der Urteile von einzelnen Lehrpersonen (zur Reliabilität und damit verbundenen Problemen vgl. Lintorf et al., 2011; Lorenz, 2011). Die Stabilität der Rangordnungskomponente bei Leistungseinschätzungen wurde von Lorenz und Artelt (2009) untersucht. Die über den Zeitraum eines halben Jahres berechneten Korrelationen (.38 für Mathematik, .51 für Textverstehen, .57 für Wortschatz;  $N = 127$  Klassen) fallen ähnlich hoch aus wie die Urteilsgüte zu den einzelnen Messzeitpunkten.

Kaum etwas ist bekannt über Bedingungsfaktoren und Korrelate der Diagnosegenauigkeit wie Berufs- bzw. Lehrerfahrung oder kognitive und nicht kognitive Persönlichkeitsmerkmale (zusammenfassend z.B. Lorenz, 2011). Einer der wenigen Befunde ist das von Karing (2009) berichtete Ergebnis, dass Grundschullehrpersonen Leistungen und Fachinteresse der Schülerinnen und Schüler genauer einschätzten als Gymnasiallehrpersonen. Dagegen hängen Unterschiede in der diagnostischen Kompetenz vom Kontext ab. Dies gilt insbesondere für die Merkmalshomogenität bzw. -heterogenität, die ein urteilerleichterndes bzw. urteilerschwerendes Kontextmerkmal für die Rangordnungskomponente darstellt, allerdings auch aus rein statistischen Gründen mit der Höhe der Korrelation zusammenhängt. Masse der Diagnosegenauigkeit sind daher keine reinen Indikatoren für Personenmerkmale, sondern eher der Urteiler-Beurteilten-Relation bzw. der Beurteiler-Umwelt-Relation (Schrader, 2009).

Ergebnisse zum Urteilsniveau (Niveauelemente) werden sehr viel seltener berichtet. Insgesamt zeigt sich hier überwiegend eine Tendenz zur Überschätzung des Leistungsniveaus (zusammenfassend van Ophuysen, 2010). So berichten etwa Bates und Nettelbeck (2001) eine Tendenz zur Überschätzung der Leseleistung von 6- bis 8-jährigen Kindern, insbesondere bei leseschwachen Schülerinnen und Schülern. Dieses Befundmuster deckt sich mit Ergebnissen der Expertiseforschung, denen zufolge Expertinnen und Experten Schwierigkeiten damit bekunden, die Leistungen von Laiinnen und Laien zutreffend einzuschätzen. In einer Studie von Herppich, Wittwer, Nückles und Renkl (2013) überschätzten Tutorinnen und Tutoren, die jeweils mit einer einzigen

oder einem einzigen Lernenden zusammenarbeiteten, deren Verständnisleistungen. Allerdings urteilten erfahrene Lehrpersonen entgegen dem allgemeinen Befund aus der Expertiseforschung genauer und konnten auch ihre eigene Diagnosefähigkeit besser einschätzen als studentische Tutorinnen und Tutoren. Bei der nur selten untersuchten Differenzierungskomponente liess sich keine einheitliche Tendenz zur Über- oder Unterschätzung feststellen (vgl. z.B. van Ophuysen, 2010). Empirische Untersuchungen (z.B. Spinath, 2005) zeigen ferner, dass die Komponenten der Diagnosegenauigkeit unabhängig sind, was die Frage aufwirft, ob bzw. in welchem Sinne man überhaupt von einem allgemeinen Konstrukt der diagnostischen Kompetenz sprechen kann.

Der Ansatz des simulierten Klassenraums, bei dem die Leistungen hypothetischer Schülerinnen und Schüler beurteilt werden, bietet die Möglichkeit, diagnostische Kompetenz unter kontrollierten Bedingungen zu untersuchen (vgl. Südkamp & Möller, 2009; Südkamp, Möller & Pohlmann, 2008). In diesem computergesteuerten System wählen die Urteilernden aus einem Menü vorgegebene Fragen aus, zu denen die von der Urteilerin oder vom Urteiler jeweils ausgewählten «Schülerinnen» und «Schüler» (gesteuert durch einen vorher eingestellten Fähigkeitsparameter und eine Zufallskomponente) eine richtige oder falsche Antwort geben, über deren Korrektheit die Probandin oder der Proband eine Rückmeldung erhält. Die nach Ende einer Unterrichtseinheit für jede Schülerin und jeden Schüler abgegebene Leistungseinschätzung (Zahl gelöster Aufgaben, Note) wird mit der tatsächlichen Leistung (Zahl der richtigen Antworten) verglichen. Bemerkenswert ist, dass Untersuchungen mit dem simulierten Klassenraum zu ähnlichen Ergebnissen führen wie im realen Klassenkontext (Rangordnungskomponente in der Höhe von .53 bis .68; leichte Tendenz zur Überschätzung des Leistungsniveaus bei einer leichten Unterschätzung der Streuung; vgl. Südkamp, Möller & Pohlmann, 2008).

### 3.3 Diagnosekompetenz, Unterrichtsqualität und Lernerfolg

Ob die Fähigkeit von Lehrpersonen, die Leistungen ihrer Schülerinnen und Schüler zutreffend zu beurteilen, auch für den Unterrichtserfolg eine Rolle spielt, ist bislang erst selten untersucht worden. Helmke und Schrader (1987) konnten keinen direkten Zusammenhang zwischen Diagnosegenauigkeit und Lernerfolg der Schülerinnen und Schüler im Fach Mathematik nachweisen. Hohe Diagnosekompetenz hatte aber dann einen Effekt, wenn Strukturierungshilfen und individuelle Hilfestellungen ausreichend häufig eingesetzt wurden (Katalysatorfunktion der diagnostischen Kompetenz; vgl. Helmke, 2012). In zwei neueren Untersuchungen konnten dagegen direkte Zusammenhänge mit dem Lernerfolg der Schülerinnen und Schüler aufgezeigt werden. In der Studie von Anders, Kunter, Brunner, Krauss und Baumert (2010) liess sich der Lernerfolg der Schülerinnen und Schüler in Mathematik durch die Genauigkeit der Schwierigkeitseinschätzung und die diagnostische Sensitivität vorhersagen, aber nur die Genauigkeit der Schwierigkeitseinschätzung hing auch mit der Unterrichtsqualität zusammen. Karing, Pfof und Artelt (2011) konnten für ein aufgabenspezifisches Treffermass einen Zusammenhang mit dem Lernerfolg im Lesen aufzeigen, nicht aber

für die Rangordnungskomponente. Das Treffermass sagte – erwartungsgemäss – den Lernerfolg im Lesen bei hoher, nicht aber bei niedriger Individualisierung sowie – erwartungskonträr – bei niedriger, nicht aber bei hoher Strukturierung vorher, während es bei der Vorhersage der Mathematikleistung für beide Diagnosemasse weder einfache noch differenzielle Effekte gab.

Das von Beck et al. (2008) entwickelte Konzept der adaptiven Lehrkompetenz, bei dem diagnostische Kompetenzen sowohl in der Planungs- wie auch in der Durchführungsphase des Unterrichts eng mit unterrichtsbezogenen Fähigkeiten (Klassenführung, Didaktik, Sachkompetenz) verknüpft sind, war Gegenstand vertiefender Analysen von Brühwiler (2012). Die adaptive Lehrkompetenz wurde in der Planungsphase mittels eines Vignetten-Ansatzes (Beschreibungen hypothetischer Situationen, zu denen die Probandinnen und Probanden Planungsüberlegungen äussern sollten) und in der Handlungsphase mittels eines Videotests (ein nach einem Drehbuch aufgezeichnetes Unterrichtsvideo, zu dem die Probandinnen und Probanden ihre Handlungen beschreiben und begründen sollten) erfasst. Mit der adaptiven Lehrkompetenz liess sich der Lernerfolg der Schülerinnen und Schüler vorhersagen. Gleichzeitig korrelierte sie in der Höhe von .39 mit der auf die Lernzielerreichung der Schülerinnen und Schüler bezogenen diagnostischen Urteilsfähigkeit (Rangordnungskomponente), was für die enge Verzahnung von diagnostischen und didaktischen Schritten spricht.

### 3.4 Diagnostische Kompetenz: Prozessorientierte Ansätze

Verschiedene Studien geben Aufschluss über die Prozesse, die diagnostischen Urteilen zugrunde liegen. Ausgehend von Befunden der Expertiseforschung (schnellere Verarbeitung, flexiblere Organisation und Nutzung von Informationen, geringere Anfälligkeit für Urteilsverzerrungen) untersuchte van Ophuysen (2006), ob sich Expertinnen und Experten (erfahrene Grundschullehrpersonen) bei der Abgabe einer Übergangsempfehlung für einen hypothetischen Schüler von Novizinnen und Novizen (Lehramtsstudierende) unterscheiden. Gezeigt werden konnte, dass Novizinnen und Novizen einer stärkeren Bestätigungstendenz (Beibehaltung des Urteils trotz inkonsistenter Informationen) unterlagen, während Expertinnen und Experten nicht nur differenzierter urteilten, sondern ihr Urteil bei inkonsistenten Zusatzinformationen auch häufiger revidierten. Untersuchungen von Krolak-Schwerdt und weiteren Kolleginnen und Kollegen (z.B. Krolak-Schwerdt & Rummer, 2005; Krolak-Schwerdt, Böhmer & Gräsel, 2009) stützen sich auf Ansätze der sozialen Urteilsforschung, bei denen zwischen einer *kategorienbasierten* (Aktivierung von Kategorien/Stereotypen) und einer *merkmalbasierten* Informationsverarbeitung (Abprüfung von Einzelmerkmalen) unterschieden wird. Krolak-Schwerdt und Rummer (2005) konnten zeigen, dass Expertinnen und Experten (Lehrpersonen) bei einer Prognoseaufgabe eine merkmalsbasierte Verarbeitung vornahmen, bei der Aufforderung, sich einen Eindruck zu bilden, dagegen kategorienbasiert vorgehen, während sich bei Laiinnen und Laien die Vorgehensweisen in beiden Fällen nicht unterschieden. Auch in einer Untersuchung von Krolak-Schwerdt, Böhmer und Gräsel (2009), in der erfahrene Lehrpersonen (Expertinnen und Exper-



ten) und Lehramtsstudierende (Novizinnen und Novizen) anhand der Fallbeschreibung eines hypothetischen Schülers entweder dessen schulische Entwicklung vorhersagen (merkmalbasierte Verarbeitung) oder sich einen Eindruck über den Schüler bilden sollten (kategorienbasierte Verarbeitung), variierten die erfahrenen Lehrpersonen ihre Urteilsbildung je nach Verarbeitungsziel und erwiesen sich damit als flexible Denkerinnen und Denker, während die Informationsverarbeitung der Laiinnen und Laien durch das Verarbeitungsziel nicht beeinflusst wurde. Krolak-Schwerdt, Böhmer und Gräsel (2013) konnten darüber hinaus zeigen, dass die Verantwortlichkeit für das Urteil (Accountability) ein Moderator für die Urteilsbildung ist. Dünnebier, Gräsel und Krolak-Schwerdt (2009) untersuchten mittels der Ankerheuristik, ob die beiden Modi der Informationsverarbeitung unterschiedlich anfällig gegenüber vereinfachenden, aber verzerrungsanfälligen heuristischen Strategien sind. Sie konnten für das Fach Deutsch, nicht aber für Mathematik aufzeigen, dass der Anker bei Expertinnen und Experten (erfahrene Lehrpersonen) unter der Bedingung einer kategorienbasierten Verarbeitung (Eindrucksbildung) einen Effekt hatte, nicht aber bei einer merkmalgeleiteten Verarbeitung (Schullaufbahnempfehlung). Bei den Novizinnen und Novizen (Lehramtsstudierende) zeigte sich unter beiden Bedingungen ein Ankereffekt.

#### 4 Fazit und Ausblick

Lehrpersonendiagnostik bewegt sich in einem Spannungsfeld zwischen formeller, d.h. auf wissenschaftlich fundierten Methoden beruhender Diagnostik und informeller Diagnostik. Informelle Lehrpersonenurteile sind unverzichtbar, wenn es um die Handlungssteuerung im Unterricht geht. Eine formelle Diagnostik ist immer dann angebracht, wenn es um Entscheidungen mit schwerwiegenden und weitreichenden Konsequenzen geht. Im Unterschied dazu können lern- und unterrichtsbezogene Entscheidungen nicht in jedem Fall durch explizite Diagnosen abgesichert werden, schon gar nicht, wenn es um die Feinabstimmung des Unterrichts (Mikroadaptation) geht.

Die Urteilsgenauigkeit wird in der Forschung als Indikator für die diagnostische Kompetenz angesehen. Untersuchungen haben gezeigt, dass die Urteilsgenauigkeit direkt oder indirekt mit dem Lernerfolg der Schülerinnen und Schüler zusammenhängt. Das Problem dieser Studien ist allerdings, dass es über die Prozesse, die zwischen den diagnostischen Urteilen der Lehrpersonen und den Lernergebnissen aufseiten der Schülerinnen und Schüler vermitteln, kaum empirisch gesichertes Wissen gibt. Dafür, dass es sich bei der diagnostischen Kompetenz um ein für die Lehrtätigkeit und insbesondere für das unterrichtliche Handeln bedeutsames Merkmal handelt, das sich auch im Gesamtgefüge der Lehrpersonenmerkmale gut verorten lässt (Baumert & Kunter, 2006), sprechen gute theoretische Gründe. Auch die Forschungsübersichten von Hattie (2009, 2012) weisen darauf hin, dass regelmässiges Feedback, eine engmaschige Überwachung des Lernerfolgs und die fortlaufende Anpassung des Unterrichtsangebots durch adaptive Lernexpertinnen und Lernexperten ganz wesentlich zum Lernerfolg der Schü-

lerinnen und Schüler beitragen. Hattie sieht darin das eigentliche Erfolgsgeheimnis für einen effektiven, leistungsförderlichen Unterricht. Die mit dem Ansatz der Urteils-genauigkeit verbundenen Probleme (ungeklärte Reliabilität, fehlende Zusammenhänge mit Lehrpersonenmerkmalen, Abhängigkeit vom Urteilskontext, Unabhängigkeit verschiedener Komponenten der Urteils-genauigkeit, unklare Wirkmechanismen; vgl. Lorenz, 2011; Schrader, 2009; Spinath, 2005) werfen allerdings Fragen auf, die im Zuge weiterer Forschung zu klären sind.

Die prozessorientierten Forschungen geben Hinweise auf Faktoren, die die Genauigkeit des Urteils beeinflussen können. Sie tragen zu einem vertieften Verständnis der Unterschiede zwischen formellen und informellen Beurteilungen und der Unterschiede in der Urteilsbildung von Expertinnen und Experten einerseits und von Novizinnen und Novizen andererseits bei. Da es sich dabei um (quasi)experimentelle Studien handelt, erlauben sie auch bessere Rückschlüsse auf die kausale Interpretierbarkeit. Um die ökologische Validität zu verbessern und fundierte Rückschlüsse auf die schulische Realsituation zu ermöglichen, wäre es sinnvoll, diese grundlagenorientierten Studien stärker mit Felduntersuchungen zu verknüpfen, insbesondere auch mit solchen zur Urteils-genauigkeit (Böhmer, Gräsel, Hörstermann & Krolak-Schwerdt, 2012; Schrader, 2009).

Vor diesem Hintergrund erscheint es sinnvoll, von einem breiteren Verständnis diagnostischer Kompetenz auszugehen als von der blossen Diagnosegenauigkeit. Dass Lehrpersonenurteile genau oder zumindest frei von groben Verfälschungen und Verzerrungen sein sollten, ist zwar mehr als naheliegend. Ebenso wichtig dürfte aber auch die Bereitschaft sein, die eigenen Urteile zu reflektieren, zu revidieren und flexibel an Veränderungen anzupassen. Zu diagnostischer und pädagogischer Professionalität gehört demnach auch die Bereitschaft, sich der eigenen, die Diagnose möglicherweise beeinträchtigenden und verzerrenden Erwartungen, Urteilstendenzen und impliziten Theorien bewusst zu werden. Eine erhöhte Sensibilität für diagnostische Fragen, eine zunehmende Einsicht in die Grenzen der eigenen Urteilsfähigkeit und die zunehmende Bereitschaft, die eigenen Urteilsleistungen durch den Einsatz von Diagnoseverfahren und den Einbezug anderer Perspektiven (z.B. Selbsteinschätzungen der Schülerinnen und Schüler oder Urteile der Kolleginnen und Kollegen) regelmässig auf den Prüfstand zu stellen und zu verbessern, könnten mindestens ebenso wichtig sein wie die Urteils-genauigkeit als solche. Zu einer professionellen Lehrpersonendiagnostik gehören zudem die Fähigkeit, geeignete Strategien und Methoden der Informationsbeschaffung und -verarbeitung einzusetzen, was auch die passgenaue Auswahl und Konstruktion von Tests und anderen Verfahren zur Lernkontrolle einschliesst, sowie die Bereitschaft, sich mit den gewonnenen Daten fundiert auseinanderzusetzen und diese sachgerecht zu interpretieren (Nitko, 2004).

Trotz der bislang noch bestehenden Forschungsdefizite werden die Förderung und die Verbesserung diagnostischer Kompetenz als wichtiges Anliegen betrachtet. Ein un-

längst im Auftrag der deutschen Kultusministerkonferenz durchgeführtes Projekt setzt dabei auf eine verstärkte Aus- und Fortbildung von Lehrpersonen in vier Teilbereichen der pädagogischen Diagnostik: Individualdiagnostik, Diagnostik im Rahmen von Vergleichsarbeiten, Diagnostik im Rahmen des Bildungsmonitorings sowie Unterrichtsdiagnostik (vgl. Bos & Hovenga, 2010). Die Wissensvermittlung erfolgt dabei durch Studienbriefe und begleitende Aufgabenstellungen und Materialien. Im Bereich der Unterrichtsdiagnostik wurden ausserdem Tools entwickelt, die nicht nur darauf abzielen, Wissen zu vermitteln, sondern auch darauf, die Urteilspraxis zu verbessern (vgl. Helmke et al., 2012) und auf diese Weise viele der oben als wünschenswert genannten Kompetenzmerkmale zu erreichen. Inwieweit dieser Ansatz zum Aufbau eines fundierten diagnostischen und metakognitiven Wissens und zur Verbesserung der diagnostischen Kompetenz beiträgt, bleibt abzuwarten. Die bislang vorliegenden Erfahrungen stimmen positiv, bedürfen aber der gründlichen Abklärung und Evaluation durch wissenschaftliche Folgeuntersuchungen.

## Literatur

- Altrichter, H.** (2010). Schul- und Unterrichtsentwicklung durch Datenrückmeldung. In H. Altrichter & K. Maag Merki (Hrsg.), *Handbuch Neue Steuerung im Schulsystem* (S. 219–254). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Anders, Y., Kunter, M., Brunner, M., Krauss, S. & Baumert, J.** (2010). Diagnostische Fähigkeiten von Mathematiklehrkräften und ihre Auswirkungen auf die Leistungen ihrer Schülerinnen und Schüler. *Psychologie in Erziehung und Unterricht*, 57 (3), 175–193.
- Artelt, C. & Gräsel, C.** (2009). Gasteditorial: Diagnostische Kompetenz von Lehrkräften. *Zeitschrift für Pädagogische Psychologie*, 23 (3–4), 157–160.
- Bates, C. & Nettelbeck, T.** (2001). Primary school teachers' judgements of reading achievement. *Educational Psychology*, 21 (2), 177–187.
- Baumert, J. & Kunter, M.** (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. *Zeitschrift für Erziehungswissenschaft*, 9 (4), 469–520.
- Beck, E., Baer, M., Guldemann, T., Bischoff, S., Brühwiler, C., Müller, P. et al.** (2008). *Adaptive Lehrkompetenz. Analyse und Struktur, Veränderbarkeit und Wirkung handlungssteuernden Lehrerwissens*. Münster: Waxmann.
- Böhmer, I., Gräsel, C., Hörstermann, T. & Krolak-Schwerdt, S.** (2012). Die Informationssuche bei der Erstellung der Übergangsempfehlung – Die Rolle von Fallkonsistenz und Expertise. *Unterrichtswissenschaft*, 40 (2), 140–155.
- Bos, W. & Hovenga, N.** (2010). Diagnostische Kompetenzen – besser individuell fördern. *Schule NRW*, Heft 8, 383–385.
- Brühwiler, C.** (2012). *Adaptive Lehrkompetenz und schulisches Lernen. Effekte handlungssteuernder Kognitionen von Lehrpersonen auf Unterrichtsprozesse und Lernergebnisse der Schülerinnen und Schüler*. Unveröffentlichte Dissertation. Landau: Universität Landau, Fachbereich Psychologie.
- Coladarsi, T.** (1986). The accuracy of teacher judgments of student responses to standardized test items. *Journal of Educational Psychology*, 78 (2), 141–146.
- Cronbach, L.J.** (1955). Processes affecting scores on «understanding of others» and «assumed similarity». *Psychological Bulletin*, 52 (3), 177–193.
- Dünnebier, K., Gräsel, C. & Krolak-Schwerdt, S.** (2009). Urteilsverzerrungen in der schulischen Leistungsbeurteilung. Eine experimentelle Studie zu Ankereffekten. *Zeitschrift für Pädagogische Psychologie*, 23 (3–4), 187–195.

- Hascher, T.** (2008). Diagnostische Kompetenzen von Lehrpersonen. In C. Kraler & M. Schratz (Hrsg.), *Wissen erwerben, Kompetenzen entwickeln. Modelle zur kompetenzorientierten Lehrerbildung* (S. 71–86). Münster: Waxmann.
- Hasselhorn, M., Schneider, W. & Marx, H.** (Hrsg.). (2000). *Diagnostik von Lese-Rechtschreibschwierigkeiten* (Jahrbuch der pädagogisch-psychologischen Diagnostik. Tests und Trends – Neue Folge, Band 1). Göttingen: Hogrefe.
- Hattie, J.** (2009). *Visible learning. A synthesis of over 800 meta-analyses relating to achievement*. London: Routledge.
- Hattie, J.** (2012). *Visible learning for teachers. Maximizing impact on learning*. London: Routledge.
- Helmke, A.** (2012). *Unterrichtsqualität und Lehrerverprofessionalität. Diagnose, Evaluation und Verbesserung des Unterrichts* (4. Auflage). Seelze: Klett-Kallmeyer.
- Helmke, A., Helmke, T., Lenske, G., Pham, G., Praetorius, A.-K., Schrader, F.-W. & Ade-Thurrow, M.** (2012). Unterrichtsdiagnostik – Voraussetzung für die Verbesserung der Unterrichtsqualität. In S.G. Huber (Hrsg.), *Jahrbuch Schulleitung 2012. Befunde und Impulse zu den Handlungsfeldern des Schulmanagements* (S. 133–144). Köln: Carl Link.
- Helmke, A. & Hosenfeld, I.** (2004). Vergleichsarbeiten – Standards – Kompetenzstufen: Begriffliche Klärung und Perspektiven. In R.S. Jäger, A. Frey & M. Wosnitza (Hrsg.), *Lernprozesse, Lernumgebungen und Lerndiagnostik* (S. 56–75). Landau: Verlag Empirische Pädagogik.
- Helmke, A. & Schrader, F.-W.** (1987). Interactional effects of instructional quality and teacher judgement accuracy on achievement. *Teaching and Teacher Education*, 3 (2), 91–98.
- Helmke, A. & Weinert, F.E.** (1997). Bedingungsfaktoren schulischer Leistungen. In F.E. Weinert (Hrsg.), *Enzyklopädie der Psychologie. Pädagogische Psychologie. Psychologie des Unterrichts und der Schule* (S. 71–176). Göttingen: Hogrefe.
- Herppich, S., Wittwer, J., Nückles, M. & Renkl, A.** (2013). Does it make a difference? Investigating the assessment accuracy of teacher tutors and student tutors. *The Journal of Experimental Education*, 81 (2), 242–260.
- Hesse, I. & Latzko, B.** (2009). *Diagnostik für Lehrkräfte*. Opladen: Barbara Budrich.
- Hoge, R.D. & Coladarsi, T.** (1989). Teacher-based judgments of academic achievement: A review of literature. *Review of Educational Research*, 59 (3), 297–313.
- Hosenfeld, I., Helmke, A. & Schrader, F.-W.** (2002). Diagnostische Kompetenz: Unterrichts- und lernrelevante Schülermerkmale und deren Einschätzung durch Lehrkräfte in der Lehr-Lern-Studie SALVE. In M. Prenzel & J. Doll (Hrsg.), *Bildungsqualität von Schule: Schulische und außerschulische Bedingungen mathematischer, naturwissenschaftlicher und überfachlicher Kompetenzen* (45. Beiheft der Zeitschrift für Pädagogik) (S. 65–82). Weinheim: Beltz.
- Ingenkamp, K. & Lissmann, U.** (2008). *Lehrbuch der Pädagogischen Diagnostik* (6. Auflage). Weinheim: Beltz.
- Karing, C.** (2009). Diagnostische Kompetenz von Grundschul- und Gymnasiallehrkräften im Leistungsbe- reich und im Bereich Interessen. *Zeitschrift für Pädagogische Psychologie*, 23 (3–4), 197–209.
- Karing, C., Matthäi, J. & Artelt, C.** (2011). Genauigkeit von Lehrerurteilen über die Lesekompetenz ihrer Schülerinnen und Schüler in der Sekundarstufe I – Eine Frage der Spezifität? *Zeitschrift für Pädagogische Psychologie*, 25 (3), 159–172.
- Karing, C., Pfost, M. & Artelt, C.** (2011). Hängt die diagnostische Kompetenz von Sekundarstufenlehrkräften mit der Entwicklung der Lesekompetenz und der mathematischen Kompetenz ihrer Schülerinnen und Schüler zusammen? *Journal for Educational Research Online*, 3 (2), 119–147.
- Karst, K.** (2012). *Kompetenzmodellierung des diagnostischen Urteils von Grundschullehrern*. Münster: Waxmann.
- 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 (3–4), 175–186.
- Krolak-Schwerdt, S., Böhmer, M. & Gräsel, C.** (2013). The impact of accountability on teachers' assessments of student performance: a social cognitive analysis. *Social Psychology of Education*, 16 (2), 215–239.

## Diagnostische Kompetenz von Lehrpersonen

- Krolak-Schwerdt, S. & Rummer, R.** (2005). Der Einfluss von Expertise auf den Prozess der schulischen Leistungsbeurteilung. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 37 (4), 205–213.
- Lintorf, K., McElvany, N., Rjosk, C., Schroeder, S., Baumert, J., Schnotz, W. et al.** (2011). Zuverlässigkeit von diagnostischen Lehrerurteilen – Reliabilität verschiedener Urteilsmaße bei der Einschätzung von Aufgabenschwierigkeiten. *Unterrichtswissenschaft*, 39 (2), 102–120.
- Lorenz, C.** (2011). *Diagnostische Kompetenz von Grundschullehrkräften. Strukturelle Aspekte und Bedingungen*. Bamberg: University of Bamberg Press.
- Lorenz, C. & Artelt, C.** (2009). Fachspezifität und Stabilität diagnostischer Kompetenz von Grundschullehrkräften in den Fächern Deutsch und Mathematik. *Zeitschrift für Pädagogische Psychologie*, 23 (3–4), 211–222.
- McElvany, N., Schroeder, S., Hachfeld, A., Baumert, J., Richter, T., Schnotz, W. et al.** (2009). Diagnostische Fähigkeiten von Lehrkräften bei der Einschätzung von Schülerleistungen und Aufgabenschwierigkeiten bei Lernmedien mit instruktionalen Bildern. *Zeitschrift für Pädagogische Psychologie*, 25 (3), 223–235.
- Nitko, A.J.** (2004). *Educational assessment of students* (4<sup>th</sup> edition). Upper Saddle River, NJ: Pearson.
- Praetorius, A.-K., Greb, K., Dickhäuser, O. & Lipowsky, F.** (2011). Wie gut schätzen Lehrer die Fähigkeitsselbstkonzepte ihrer Schüler ein? Zur diagnostischen Kompetenz von Lehrkräften. *Psychologie in Erziehung und Unterricht*, 58 (2), 81–91.
- Schrader, F.-W.** (2009). Anmerkungen zum Themenschwerpunkt Diagnostische Kompetenz von Lehrkräften. *Zeitschrift für Pädagogische Psychologie*, 23 (3–4), 237–245.
- Schrader, F.-W.** (2011). Lehrer als Diagnostiker. In E. Terhart, H. Bennewitz & M. Rothland (Hrsg.), *Handbuch der Forschung zum Lehrerberuf* (S. 683–698). Münster: Waxmann.
- Schrader, F.-W. & Helmke, A.** (1987). Diagnostische Kompetenz von Lehrern: Komponenten und Wirkungen. *Empirische Pädagogik*, 1 (1), 27–52.
- Schrader, F.-W. & Helmke, A.** (2001). Alltägliche Leistungsbeurteilung durch Lehrer. In F.E. Weinert (Hrsg.), *Leistungsmessungen in Schulen* (S. 45–58). Weinheim: Beltz.
- Shepard, L.A.** (2001). The role of classroom assessment in teaching and learning. In V. Richardson (Hrsg.), *Handbook of research on teaching* (4<sup>th</sup> edition) (S. 1066–1101). Washington, DC: American Educational Research Association.
- Spinath, B.** (2005). Akkurathheit der Einschätzung von Schülermerkmalen durch Lehrer/innen und das Konstrukt der diagnostischen Kompetenz. *Zeitschrift für Pädagogische Psychologie*, 19 (1–2), 85–95.
- Südkamp, A., Kaiser, J. & Möller, J.** (2012). Accuracy of teachers' judgments of students' academic achievement: A meta-analysis. *Journal of Educational Psychology*, 104 (3), 743–762.
- Südkamp, A. & Möller, J.** (2009). Referenzgruppeneffekte im Simulierten Klassenraum: direkte und indirekte Einschätzungen von Schülerleistungen. *Zeitschrift für Pädagogische Psychologie*, 23 (3–4), 161–174.
- Südkamp, A., Möller, J. & Pohlmann, B.** (2008). Der Simulierte Klassenraum. Eine experimentelle Untersuchung zur diagnostischen Kompetenz. *Zeitschrift für Pädagogische Psychologie*, 22 (3–4), 261–276.
- van Ophuysen, S.** (2006). Vergleich diagnostischer Entscheidungen von Novizen und Experten am Beispiel der Schullaufbahneempfehlung. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 38 (4), 154–161.
- van Ophuysen, S.** (2010). Professionelle pädagogisch-diagnostische Kompetenz – eine theoretische und empirische Annäherung. In N. Berkemeyer, W. Bos, H.G. Holtappels, N. McElvany & R. Schulz-Zander (Hrsg.), *Jahrbuch der Schulentwicklung. Band 16: Daten, Beispiele und Perspektiven* (S. 203–234). Weinheim: Juventa.

## Autor

**Friedrich-Wilhelm Schrader**, Dr., Universität Koblenz-Landau, Campus Landau, Fachbereich Psychologie, schrader@uni-landau.de