

Pädagogische Psychologie und  
Entwicklungspsychologie

herausgegeben von D. H. Rost

94

A. Südkamp, A.-K. Praetorius (Hrsg.)

# Diagnostische Kompetenz von Lehrkräften

WAXMANN

# **Pädagogische Psychologie und Entwicklungspsychologie**

herausgegeben von Detlef H. Rost

## **Wissenschaftlicher Beirat**

Jürgen Baumert (Berlin)  
Oliver Dickhäuser (Mannheim)  
Marcus Hasselhorn (Frankfurt)  
Andreas Knapp (Wildbad)  
Olaf Köller (Kiel)  
Detlev Leutner (Essen)  
Sabina Pauen (Heidelberg)  
Ulrich Schiefele (Potsdam)  
Christiane Spiel (Wien)  
Sabine Weinert (Bamberg)

## **Editorial**

Pädagogische Psychologie und Entwicklungspsychologie sind seit jeher zwei miteinander eng verzahnte Teildisziplinen der Psychologie. Beide haben einen festen Platz im Rahmen der Psychologenausbildung: Pädagogische Psychologie als wichtiges Anwendungsfach im zweiten Studienabschnitt, Entwicklungspsychologie als bedeutsames Grundlagenfach in der ersten und als Forschungsvertiefung in der zweiten Studienphase. Neue Zielsetzungen, neue thematische Schwerpunkte und Fragestellungen sowie umfassendere Forschungsansätze und ein erweitertes Methodenspektrum haben zu einer weiteren Annäherung beider Fächer geführt und sie nicht nur für Studierende, sondern auch für die wissenschaftliche Forschung zunehmend attraktiver werden lassen. „Pädagogische Psychologie und Entwicklungspsychologie“ nimmt dies auf, fördert die Rezeption einschlägiger guter und interessanter Forschungsarbeiten, stimuliert die theoretische, empirische und methodische Entfaltung beider Fächer und gibt fruchtbare Impulse zu ihrer Weiterentwicklung einerseits und zu ihrer gegenseitigen Annäherung andererseits.

Der Beirat der Reihe „Pädagogische Psychologie und Entwicklungspsychologie“ repräsentiert ein breites Spektrum entwicklungspsychologischen und pädagogisch-psychologischen Denkens und setzt Akzente, indem er auf Forschungsarbeiten aufmerksam macht, die den wissenschaftlichen Diskussionsprozess beleben können. Es ist selbstverständlich, dass zur Sicherung des Qualitätsstandards dieser Reihe jedes Manuskript – wie bei Begutachtungsverfahren in anerkannten wissenschaftlichen Zeitschriften – einem Auswahlverfahren unterzogen wird („peer review“). Nur qualitätsvolle Arbeiten werden der zunehmenden Bedeutung der Pädagogischen Psychologie und Entwicklungspsychologie für die Sozialisation und Lebensbewältigung von Individuen und Gruppen in einer immer komplexer werdenden Umwelt gerecht.

Anna Südkamp, Anna-Katharina Praetorius (Hrsg.)

# Diagnostische Kompetenz von Lehrkräften

Theoretische und methodische  
Weiterentwicklungen



Waxmann 2017  
Münster • New York

### **Bibliografische Informationen der Deutschen Nationalbibliothek**

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.dnb.de> abrufbar.

### **Pädagogische Psychologie und Entwicklungspsychologie; Bd. 94**

herausgegeben von Prof. Dr. Detlef H. Rost  
Philipps-Universität Marburg

Fon: 0 64 21 / 2 82 17 27

Fax: 0 64 21 / 2 82 39 10

E-Mail: [rost@mail.uni-marburg.de](mailto:rost@mail.uni-marburg.de)

ISSN 1430-2977

Print-ISBN 978-3-8309-3596-4

E-Book-ISBN 978-3-8309-8596-9

© Waxmann Verlag GmbH, 2017

[www.waxmann.com](http://www.waxmann.com)

[info@waxmann.com](mailto:info@waxmann.com)

Umschlaggestaltung: Pleßmann Design, Ascheberg

Satz: Stoddart Satz & Layout, Münster

Druck: Hubert & Co., Göttingen

Gedruckt auf alterungsbeständigem Papier,  
säurefrei gemäß ISO 9706



Printed in Germany

Alle Rechte vorbehalten. Nachdruck, auch auszugsweise, verboten.

Kein Teil dieses Werkes darf ohne schriftliche Genehmigung des Verlages in irgendeiner Form reproduziert oder unter Verwendung elektronischer Systeme verarbeitet, vervielfältigt oder verbreitet werden.

# Inhalt

<i>Anna Südkamp &amp; Anna-Katharina Praetorius</i>	
Editorial.....	11
<i>Anna-Katharina Praetorius &amp; Anna Südkamp</i>	
1 Eine Einführung in das Thema der diagnostischen Kompetenz von Lehrkräften.....	13
1.1 Diagnostische Kompetenz von Lehrkräften: Ein relevantes Thema?.....	13
1.2 Desiderata in der empirischen Erforschung diagnostischer Kompetenz.....	14
1.2.1 Definition und Konzeptualisierung .....	14
1.2.2 Erfassungs- und Analysemethoden .....	15
1.2.3 Ansätze zur Förderung.....	16
1.2.4 Empirische Belege für die Relevanz diagnostischer Kompetenz.....	16
1.3 Die Idee und Struktur des vorliegenden Bandes.....	17
<i>Karina Karst &amp; Natalie Förster</i>	
2 Ansätze zur Modellierung diagnostischer Kompetenz.....	19
<i>Karina Karst</i>	
2.1 Akkurate Urteile – die Ansätze von Schrader (1989) und McElvany et al. (2009).....	21
2.1.1 Das Modell nach Schrader (1989).....	21
2.1.2 Die Erweiterung von McElvany et al. (2009).....	23
2.1.3 Empirische Befunde.....	23
2.1.4 Stärken und Schwächen.....	24
<i>Karina Karst</i>	
2.2 Diagnostische Kompetenz und unterrichtliche Situationen.....	25
2.2.1 Empirische Befunde.....	27
2.2.2 Stärken und Schwächen.....	28
<i>Annika Ohle, Nele McElvany, Holger Horz &amp; Mark Ullrich</i>	
2.3 Aspekte diagnostischer Kompetenz im Projekt „BiTe“.....	29
2.3.1 Empirische Befunde.....	30
2.3.2 Schwächen und Stärken.....	33
<i>Anna Südkamp, Johanna Kaiser &amp; Jens Möller</i>	
2.4 Ein heuristisches Modell der Akkuratheit diagnostischer Urteile von Lehrkräften.....	33
2.4.1 Akkuratheit der Einschätzung von Schülerleistungen durch Lehrkräfte....	34
2.4.2 Zusammenhänge zwischen Lehrermerkmalen und Urteilsakkuratheit.....	35

## 6 Inhalt

2.4.3	Zusammenhänge zwischen Urteilsmerkmalen und Urteilsakkuratheit .....	35
2.4.4	Zusammenhänge zwischen Schülermerkmalen und Urteilsakkuratheit .....	36
2.4.5	Zusammenhänge zwischen Testmerkmalen und Urteilsakkuratheit .....	37
2.4.6	Fazit .....	37

### *Lars Behrmann & Stefanie van Ophuysen*

2.5	Das Vier-Komponenten-Modell der Diagnosequalität .....	38
2.5.1	Prozesskomponenten der Diagnostik .....	40
2.5.2	Ergebniskomponenten der Diagnostik .....	40
2.5.3	Diagnostische Kompetenz im Sinne des Vier-Komponenten-Modells .....	41
2.5.4	Empirische Fundierung des Vier-Komponenten-Modells .....	41
2.5.5	Besondere Aspekte des Vier-Komponenten-Modells .....	41

### *Inga Glogger-Frey & Stephanie Herppich*

2.6	Formative Diagnostik als Teilaspekt diagnostischer Kompetenz .....	42
2.6.1	Begriff und Ziel der formativen Diagnostik .....	43
2.6.2	Arten und Methoden von formativer Diagnostik .....	44
2.6.3	Aspekte formativer Diagnostik in einem neuen Kompetenzmodell .....	45

### *Natalie Förster & Ines Böhmer*

2.7	Das Linsenmodell – Grundlagen und exemplarische Anwendungen in der pädagogisch-psychologischen Diagnostik .....	46
2.7.1	Das Linsenmodell .....	47
2.7.2	Welche Moderatoren beeinflussen die Urteilsakkuratheit? .....	49
2.7.3	Anwendung des Linsenmodells in der pädagogisch-psychologischen Diagnostik .....	49
2.7.4	Fazit .....	50

### *Matthias Böhmer, Birte English & Ines Böhmer*

2.8	Schülerbeurteilungen aus der Perspektive dualer Prozessmodelle der sozialen Urteilsbildung .....	50
-----	---	----

### *Julia Klug*

2.9	Ein Prozessmodell zur Diagnostik und Förderung von selbstreguliertem Lernen .....	54
2.9.1	Fokus des Prozessmodells .....	55
2.9.2	Inhalte des Prozessmodells .....	55
2.9.3	Empirische Befundlage zum Modell .....	56
2.9.4	Kritische Würdigung des Modells .....	58

### *Lars Behrmann & Johanna Kaiser*

2.10	Das Modell pädagogischer Diagnostik nach Ingenkamp und Lissmann .....	59
------	---	----

### *Natalie Förster & Karina Karst*

2.11	Modelle diagnostischer Kompetenz: Gemeinsamkeiten und Unterschiede .....	63
------	---	----

*Stefan Ufer & Detlev Leutner*

3	Kompetenzen als Dispositionen – Begriffsklärungen und Herausforderungen.....	67
3.1	Kompetenzbegriff und Kompetenzmodelle.....	67
3.2	Herausforderungen bei der Konzeptualisierung und Modellierung von Kompetenzen .....	69
3.3	Ausblick.....	73

*Stephanie Herppich, Anna-Katharina Praetorius, Andreas Hetmanek,  
Inga Glogger-Frey, Stefan Ufer, Detlev Leutner, Lars Behrmann, Ines Böhmer,  
Matthias Böhmer, Natalie Förster, Johanna Kaiser, Constance Karing, Karina  
Karst, Julia Klug, Annika Ohle & Anna Südkamp*

4	Ein Arbeitsmodell für die empirische Erforschung der diagnostischen Kompetenz von Lehrkräften .....	75
4.1	Einführung .....	75
4.2	Ziel der Modellentwicklung.....	75
4.3	Theoretische Einbettung .....	76
4.3.1	Gegenstandsbereich .....	76
4.3.2	Kompetenzdefinition.....	78
4.4	Elemente des Arbeitsmodells.....	80
4.4.1	Kognitive Leistungsdisposition .....	82
4.4.2	Sonstige Leistungsdispositionen.....	83
4.4.3	Situationen .....	84
4.4.4	Diagnostik.....	85
4.4.5	Performanz.....	88
4.5	Bezug zu und Abgrenzung von vorhandenen Ansätzen der Modellierung diagnostischer Kompetenz .....	88
4.5.1	Integration vorhandener Ansätze .....	88
4.5.2	Abgrenzung des Arbeitsmodells von anderen Ansätzen.....	90
4.6	Zusammenfassung .....	91

*Anna-Katharina Praetorius, Andreas Hetmanek, Stephanie Herppich  
& Stefan Ufer*

5	Herausforderungen bei der empirischen Erforschung diagnostischer Kompetenz.....	95
---	--	----

*Karina Karst, Julia Klug & Stefan Ufer*

5.1	Strukturierung diagnostischer Situationen im inner- und außerunterrichtlichen Handeln von Lehrkräften.....	102
5.1.1	Gemeinsame Definitionsmerkmale diagnostischer Situationen .....	103
5.1.2	Randbedingungen .....	104
5.1.3	Gründe für die Variabilität – Klassifikationsmerkmale .....	105
5.1.4	Lehrerbefragung zur Validierung der Klassifizierungsmerkmale diagnostischer Situationen.....	111
5.1.5	Zusammenfassung .....	113

*Johanna Kaiser, Anna-Katharina Praetorius, Anna Südkamp & Stefan Ufer*

5.2 Die enge Verwobenheit von diagnostischem und pädagogischem Handeln als Herausforderung bei der Erfassung diagnostischer Kompetenz..... 114

5.2.1 Einordnung bisheriger Erfassungen diagnostischer Kompetenz ..... 116

5.2.2 Fazit und Ausblick ..... 121

*Christian Brühwiler*

5.3 Diagnostische und didaktische Kompetenz als Kern adaptiver Lehrkompetenz ..... 123

5.3.1 Adaptive Lehrkompetenz..... 124

5.3.2 Fragestellungen..... 126

5.3.3 Methodisches Vorgehen ..... 126

5.3.4 Ergebnisse ..... 130

5.3.5 Diskussion..... 132

*Lars Behrmann & Inga Glogger-Frey*

5.4 Produkt- und Prozessindikatoren diagnostischer Kompetenz ..... 134

5.4.1 Produktindikatoren diagnostischer Kompetenz ..... 134

5.4.2 Defizite der Urteilsakkurtheit..... 135

5.4.3 Prozessindikatoren der diagnostischen Kompetenz ..... 137

5.4.4 Fazit ..... 141

*Anna F. Altmann & Matthias Nückles*

5.5 Empirische Studie zu Qualitätsindikatoren für den diagnostischen Prozess ..... 142

5.5.1 Prozessorientierte Diagnose in Dialogen mit Schülerinnen und Schülern ..... 142

5.5.2 Prozessorientierte Diagnose im Dialog fördern..... 145

5.5.3 Bewertung der Qualitätsindikatoren für den diagnostischen Prozess..... 148

5.5.4 Fazit ..... 149

*Anna Südkamp & Anna-Katharina Praetorius*

6 Innovative Ansätze zur Analyse der Urteilsakkurtheit von Lehrkräften..... 151

*Karina Karst, Johannes Hartig, Johanna Kaiser & Frank Lipowsky*

6.1 Mehrebenenmodelle als Werkzeuge zur Analyse diagnostischer Kompetenz von Lehrkräften – ein lineares Mischmodell (LMM) und seine Anwendung in R..... 153

6.1.1 Kritische Würdigung bisheriger Analyseverfahren ..... 153

6.1.2 Lineare Mischmodelle (LMM) ..... 156

6.1.3 Anwendung des LMM zur Schätzung von Kennwerten diagnostischer Kompetenz von Lehrkräften ..... 161

6.1.4 Abschluss und Zusammenfassung..... 173

*Sarah Humberg, Natalie Förster, Johanna Kaiser &  
Felix D. Schönbrodt*

6.2	Konsequenzen akkurater Lehrerurteile: Response-Surface-Analyse als statistisches Verfahren zur Untersuchung von Übereinstimmungshypothesen.....	174
6.2.1	Anwendungsgebiete der RSA.....	175
6.2.2	Der intuitive Ansatz zur Testung von Übereinstimmungshypothesen .....	176
6.2.3	Übereinstimmungshypothesen mit der RSA testen .....	178
6.2.4	Weitere Hypothesen mit der RSA testen.....	188
6.2.5	Diskussion.....	199

*Constance Karing & Tina Seidel*

7	Förderung diagnostischer Kompetenz .....	201
---	--	-----

*Stephanie Herppich, Anna F. Altmann, Jörg Wittwer,  
Matthias Nückles*

7.1	Förderung von Instruktionsstrategien zum verbesserten Diagnostizieren im Eins-zu-Eins-Tutoring.....	203
7.1.1	Charakterisierung des Tutorings und der dafür bedeutsamen Aspekte diagnostischer Kompetenz.....	203
7.1.2	Instruktionale Ansätze zur Förderung des Diagnostizierens im Tutoringprozess.....	204
7.1.3	Ausblick.....	208

*Andreas Hetmanek & Tamara van Gog*

7.2	Förderung von diagnostischer Kompetenz: Potenziale von Ansätzen aus der medizinischen Ausbildung.....	209
7.2.1	Zwei Modi des Diagnostizierens und entsprechende Förderansätze.....	209
7.2.2	Charakterisierung der Typ 1- und Typ 2-Prozesse.....	210
7.2.3	Theoriegeleitete Entwicklung optimierter Trainingsansätze .....	211
7.2.4	Zusammenfassung und Ausblick.....	215

*Inga Glogger-Frey & Alexander Renkl*

7.3	Diagnostische Kompetenz fördern – Vorwissen aufgreifende Methoden in Kombination mit beispielbasiertem Kurztraining .....	217
7.3.1	Vorwissen aufgreifende Methoden und beispielbasiertes Training .....	218
7.3.2	Haben sich die Förderansätze empirisch bewährt?.....	221
7.3.3	Ausblick.....	222

*Julia Klug*

7.4	Tagebücher zum Diagnostizieren und zur Förderung von Diagnostik und Lernerfolg .....	223
7.4.1	Tagebücher.....	224
7.4.2	Empirische Ergebnisse zur Förderung diagnostischer Kompetenz von Lehrkräften durch Tagebücher.....	227

## 10 Inhalt

7.4.3	Einsatz von Tagebüchern zur Diagnostik und Förderung lernrelevanter Merkmale bei Schülerinnen und Schülern.....	230
7.4.4	Ausblick.....	230
<i>Natalie Förster &amp; Elmar Souvignier</i>		
7.5	Förderung diagnostischer Kompetenz durch Bereitstellung formativer Diagnostik.....	231
7.5.1	Formative Diagnostik von Lernverläufen.....	232
7.5.2	Forschungsbefunde zur Bereitstellung diagnostischer Informationen über Lernverläufe.....	233
7.5.3	Zusammenfassung und Ausblick.....	238
<i>Constance Karing &amp; Tina Seidel</i>		
7.6	Ausblick zur Förderung diagnostischer Kompetenz.....	240
7.6.1	Welche Ansatzpunkte gibt es zur Förderung der diagnostischen Kompetenz?.....	240
7.6.2	Wie sollten Fördermaßnahmen zur diagnostischen Kompetenz gestaltet werden?.....	241
7.6.3	Fazit.....	245
<i>Friedrich-Wilhelm Schrader</i>		
8	Diagnostische Kompetenz von Lehrkräften – Anmerkungen zur Weiterentwicklung des Konstrukts.....	247
8.1	Modellierung diagnostischer Kompetenz.....	248
8.2	Diagnose als Produkt.....	249
8.3	Diagnostischer Prozess.....	250
8.4	Diagnostik und pädagogisches Handeln.....	251
8.5	Förderung und Entwicklung diagnostischer Kompetenz.....	253
8.6	Fazit und Ausblick.....	254
Literatur.....		257
Verzeichnis der Autorinnen und Autoren.....		283

## Editorial

*Anna Südkamp & Anna-Katharina Praetorius*

Diagnostische Kompetenz von Lehrkräften – so finden wir – ist ein wichtiges, spannendes und lebendiges Forschungsfeld. Die hohe praktische Relevanz des Themas ist darauf zurückzuführen, dass Lehrkräfte ihren Unterricht ohne adäquate Kenntnisse der Leistungen und Lernvoraussetzungen ihrer Schülerinnen und Schüler nicht adaptiv gestalten können (Artelt & Drechsel, 2014; Retelsdorf & Möller, 2016). Spannend ist das Thema „diagnostische Kompetenz“ zum einen, weil es noch viele unerforschte Aspekte gibt, zum anderen aber auch, weil die bisherigen Befunde zum Teil nicht konsistent oder unerwartet sind. Die Lebendigkeit des Forschungsfelds schließlich zeigt sich u.a. darin, dass national und international regelmäßig neue Forschungsarbeiten in den Erziehungswissenschaften, den Fachdidaktiken und in der Psychologie zu dieser Thematik entstehen.

Im Austausch über unsere eigene Forschung entwickelten wir zu Beginn des Jahres 2013 die Idee, ein wissenschaftliches Netzwerk zur diagnostischen Kompetenz von Lehrkräften zu gründen. Ziel sollte es sein, die Forschungsaktivitäten in diesem Bereich zu bündeln und neue Forschungsideen zu erarbeiten. Seit Beginn des Jahres 2015 wurde das wissenschaftliche Netzwerk zur diagnostischen Kompetenz von Lehrkräften (NeDiKo) dann durch die Deutsche Forschungsgemeinschaft (DFG) gefördert. In diesem Rahmen werden bis zu fünf mehrtägige Arbeitstreffen finanziert. Die DFG ermöglicht es außerdem, Expertinnen und Experten zu den Arbeitstreffen einzuladen.

Dem Netzwerk gehören 14 Nachwuchswissenschaftlerinnen und -wissenschaftler an, die zur diagnostischen Kompetenz von Lehrkräften forschen und die sich seit Beginn seiner Initiierung aktiv in die Arbeit des Netzwerks einbringen (für eine namentliche Nennung siehe das Autorenverzeichnis am Ende des Buchs). Im Rahmen der zweijährigen Laufzeit des NeDiKo fanden vier Arbeitstreffen mit einer Dauer zwischen zwei und fünf Tagen statt. Bei jedem Arbeitstreffen befassten wir uns mit einem von vier thematischen Schwerpunkten mit Bezug zur diagnostischen Kompetenz (Modellierung, Erfassung, Analyse und Förderung), die wir im Zuge der Beantragung des Netzwerks als Lücken im Forschungsfeld identifiziert hatten. Wir hatten das Glück, für jedes der Arbeitstreffen herausragende Expertinnen und Experten gewinnen zu können, die sich zudem auch über die Arbeitstreffen hinaus in die Netzwerkarbeit eingebracht haben.

An dieser Stelle möchten wir aufgrund unserer überaus positiven Erfahrungen in den letzten zwei Jahren mit dem Netzwerk die Gelegenheit nutzen, uns herzlich zu bedanken. Ein großer Dank geht an alle Netzwerkmitglie-

der, die ihre Kompetenz in dieses Netzwerk eingebracht haben. Mit Bezug auf den Kompetenzbegriff können wir eindeutig sagen, dass es die zahlreichen Teilkompetenzen der beteiligten Personen waren, die zum Erfolg des Netzwerks beigetragen haben. Dazu zählen insbesondere die Fähigkeiten, die Arbeitstreffen vorzubereiten und zu organisieren, leidenschaftlich zu diskutieren, eigene Ideen einzubringen und offen zu teilen, sich zuverlässig an Absprachen zu halten und mit großem Engagement die Ziele des Netzwerks zu verfolgen. Ein ebenso großer Dank geht an die Expertinnen und Experten, die das NeDiKo vielfältig unterstützt haben – sei es durch spannende Vorträge im Rahmen der Arbeitstreffen, durch Workshops zu spezifischen Themen, durch ihr Einbringen in die inhaltliche Diskussion, durch Mentoring der Nachwuchswissenschaftlerinnen und -wissenschaftler oder durch ihre Mitarbeit an diesem Buch. Zwei unserer Experten nahmen ihre Rolle gar so ernst, dass wir sie mittlerweile eher als Mitglieder des Netzwerks bezeichnen würden – vielen Dank dafür, Stefan Ufer und Detlev Leutner. Ein herzliches Dankeschön geht aber auch an alle anderen Personen, die zum Gelingen dieses Buchs beigetragen haben. Das sind neben weiteren Autorinnen und Autoren insbesondere Detlef Rost und Beate Plugge, die unser Vorhaben von Beginn an positiv unterstützt haben, sowie die beiden anonymen Gutachter/innen des Buchs. Schließlich danken wir Friedrich-Wilhelm Schrader sehr herzlich, der seit Langem und bis heute (siehe Schrader, 1989; 2013) zur diagnostischen Kompetenz forscht und sich bereit erklärt hat, die Inhalte dieses Buchs zu diskutieren.

Dieses Buch ist im Wesentlichen eine Zusammenfassung der Ergebnisse unserer Netzwerkarbeit, in der wir uns wissenschaftlich mit der Modellierung, Erfassung, Analyse und Förderung diagnostischer Kompetenz auseinandergesetzt haben. Es richtet sich daher vornehmlich an Personen, die wissenschaftlich zu diesem Thema arbeiten. Das Buch liefert Informationen zu verschiedenen Teilaspekten diagnostischer Kompetenz und eignet sich daher auch als Nachschlagewerk. Wir denken, dass von der Nutzung als Nachschlagewerk auch interessierte Studentinnen und Studenten, Lehrkräfte oder im Bildungsmanagement tätige Personen profitieren können. In jedem Fall hoffen wir, dass dieses Buch denjenigen, die sich aktuell und in Zukunft für das Thema diagnostische Kompetenz interessieren, nützliche Anregungen liefert.

# **1 Eine Einführung in das Thema der diagnostischen Kompetenz von Lehrkräften**

*Anna-Katharina Praetorius & Anna Südkamp*

## **1.1 Diagnostische Kompetenz von Lehrkräften: Ein relevantes Thema?**

Lehrerkompetenzen stehen seit einigen Jahren verstärkt im Fokus der empirischen Bildungsforschung. Eine grundlegende Systematisierung der unterschiedlichen Konstrukte, die hierzu gezählt werden, wurde 2006 von Baumert und Kunter vorgelegt. Ein Schwerpunkt ihres Beitrags ist das Professionswissen von Lehrkräften. Dieses unterteilen sie in pädagogisches Wissen, Fachwissen, fachdidaktisches Wissen, Organisationswissen und Beratungswissen. Diagnostische Kompetenz lässt sich laut Baumert und Kunter (2006) in den Bereich pädagogischen Wissens und Könnens einordnen, die Autoren weisen jedoch auch auf fachdidaktische Wissensanteile der diagnostischen Kompetenz hin. Diagnostische Fähigkeiten zählen den Autoren zufolge zu den aktuell am stärksten diskutierten Aspekten der professionellen Kompetenz von Lehrkräften.

Begründet wird die Bedeutung diagnostischer Kompetenz meist mit dem Ziel der Unterrichtsadaptivität (z.B. Beck et al., 2008; Bruder, Klug, Hertel & Schmitz, 2010; Lorenz & Artelt, 2009; Praetorius, Greb, Lipowsky & Gollwitzer, 2010; Rogalla & Vogt, 2008; Südkamp, Möller & Pohlmann, 2008): Unterricht sollte idealerweise an die Lernvoraussetzungen der Schülerinnen und Schüler angepasst sein, sollte die Schülerinnen und Schüler also weder über- noch unterfordern und auch motivationale sowie emotionale Schülermerkmale berücksichtigen. In der Forschung zu diagnostischer Kompetenz von Lehrkräften wird argumentiert, dass Unterricht nur dann erfolgreich sein kann, wenn eine solche Anpassung erfolgt. Beispielsweise sollten Lehrkräfte ihren Unterricht abhängig davon planen, wie sie verschiedene Merkmale der Schülerinnen und Schüler einer Klasse (z.B. deren Leistungsfähigkeit oder Motivation) beurteilen (Alvidrez & Weinstein, 1999). Diese Einschätzungen können u.a. die Auswahl von Unterrichtsaktivitäten, von Aufgaben mit unterschiedlichem Schwierigkeitsgrad, von Teststrategien und die Entscheidung für bestimmte Arbeitsformen beeinflussen. Die theoretisch angenommene hohe Bedeutung diagnostischer Kompetenz für die Lernwirksamkeit von Unterricht lässt sich daraus ableiten, dass sich alle genannten Anpassungen nur dann positiv auswirken können, wenn die zugrunde liegenden Schülerdiagnosen zutreffend sind (Helmke, 2009; Weinert, Schrader & Helmke, 1990).

## **1.2 Desiderata in der empirischen Erforschung diagnostischer Kompetenz**

Die hohe Relevanz der diagnostischen Kompetenz als Teilaspekt der professionellen Kompetenz von Lehrkräften schlägt sich in einem anhaltenden Forschungsinteresse an diesem Thema nieder. So sind in den letzten Jahren zahlreiche Forschungsarbeiten entstanden, die vielfältige Erkenntnisse über die diagnostische Kompetenz von Lehrkräften hervorgebracht haben, jedoch zeitgleich auch auf die mit der Erforschung diagnostischer Kompetenz verbundenen Schwierigkeiten verweisen. Auf die aus unserer Sicht zentralen Herausforderungen in den Bereichen der Definition und Konzeptualisierung, der Methoden zur Erfassung und Analyse sowie der Förderung diagnostischer Kompetenz wird im Folgenden eingegangen.

### **1.2.1 Definition und Konzeptualisierung**

Schrader (2013) zufolge bezieht sich diagnostische Kompetenz auf „die Fähigkeit, die im Lehrberuf anfallenden diagnostischen Aufgabenstellungen erfolgreich zu bewältigen, und auf die Qualität der dabei erbrachten Diagnoseleistungen“ (S. 154). In empirischen Forschungsarbeiten wird in der Regel eine engere Definition gewählt und diagnostische Kompetenz als die Fähigkeit beschrieben, Schüler- und Aufgabenmerkmale korrekt zu beurteilen. Im vorliegenden Band wird in Bezug auf die Akkuratheit von Lehrereinschätzungen von Urteilsakkuratheit gesprochen, in Anlehnung an den entsprechenden englischen Begriff „judgment accuracy“ (siehe auch Spinath, 2005). Auf den im deutschsprachigen Raum weitverbreiteten Begriff der Urteilsgenauigkeit wird verzichtet, da der Begriff der Genauigkeit mit der Reliabilität von Maßen assoziiert ist, die jedoch nicht im Fokus der Forschung zur Urteilsakkuratheit steht. Hier wird vielmehr die Übereinstimmung von Lehrereinschätzungen und über Tests bzw. Fragebögen gemessenen Merkmalsausprägungen in den Blick genommen. Helmke, Hosenfeld und Schrader (2004) bezeichnen diese Übereinstimmung auch als Veridikalität, die als eine Spezialform der Validität gilt. Bei der Veridikalität muss laut Helmke et al. (2004) berücksichtigt werden, dass hier davon ausgegangen wird, dass Prädiktor und Kriterium das gleiche Merkmal (z.B. Schulleistungen) erfassen. Im Vergleich dazu werden bei der kriterienbezogenen Validität Beziehungen zwischen ähnlichen, aber nicht unbedingt gleichen Merkmalen beschrieben. Neben dem Forschungsstrang, in dem die Urteilsakkuratheit von Lehrkräften untersucht wird, hat sich in den letzten Jahren eine Forschungsrichtung entwickelt, die

sich anhand von Modellen der sozialen Kognition mit den Informationsverarbeitungsprozessen beschäftigt, die Lehrerdiagnosen zugrunde liegen (z.B. Krolak-Schwerdt, Böhmer & Gräsel, 2009).

An dem bisherigen Verständnis diagnostischer Kompetenz werden einige Aspekte kritisiert. So wird zunächst einmal darauf verwiesen, dass die Forschung zur Urteilsakkuratheit und die prozessorientierte Forschung bislang unverbunden nebeneinanderstehen (Schrader, 2009). Zudem wird die Gleichsetzung von diagnostischer Kompetenz und Urteilsakkuratheit in den letzten Jahren vermehrt infrage gestellt (z.B. Bruder et al., 2010; Helmke, 2009; van Ophuysen, 2010). Abs (2007) merkt beispielsweise an, dass die didaktische Relevanz akkurater Diagnosen unklar ist, und van Ophuysen (2010, S. 223) schreibt, dass „reine Ergebnisqualität einer Diagnose kein Beleg für ihre Professionalität ist“. In diesem Zusammenhang wurden auch erste theoretische Modelle zur diagnostischen Kompetenz von Lehrkräften entwickelt, die über die Urteilsakkuratheit von Lehrkräften hinausgehen (z.B. Klug, Bruder, Kelava, Spiel & Schmitz, 2013; van Ophuysen, 2010). Es fehlt jedoch ein umfassendes und überprüfbares theoretisches Rahmenmodell der diagnostischen Kompetenz von Lehrkräften, das unterschiedliche Aspekte diagnostischer Kompetenz (z.B. diagnostisches Wissen, Diagnoseanlässe, Konsequenzen des diagnostischen Urteils und den Diagnoseprozess) integriert.

## 1.2.2 Erfassungs- und Analysemethoden

Die Gleichsetzung von diagnostischer Kompetenz und Urteilsakkuratheit schlägt sich zudem oft auch in einem eher einseitigen Einsatz von Forschungsmethoden nieder. So werden in den meisten Studien Lehrereinschätzungen zu einem einzelnen oder einigen wenigen Schülermerkmalen über Tests bzw. Fragebögen erhoben (einen Überblick für den Leistungsbereich liefern z.B. Südkamp, Kaiser & Möller, 2012). Hier ist es zum einen notwendig, die verwendeten Verfahren zu optimieren, und zum anderen, neuartige Zugänge zur Erfassung der diagnostischen Kompetenz von Lehrkräften zu wählen.

Ein weiterer Kritikpunkt bezieht sich auf die Methoden zur Analyse der Urteilsakkuratheit. Als Maß für die Urteilsakkuratheit werden in der deutschsprachigen Literatur v.a. die drei von Schrader und Helmke (1987) vorgeschlagenen Akkuratheitskomponenten verwendet. Trotz ihrer hohen Popularität sind die Urteilskomponenten mit einigen Problemen behaftet. So erlauben die Kennwerte beispielsweise lediglich Aussagen auf Gruppenebene, zudem existieren bislang keine eindeutigen Richtwerte zu ihrer Interpretation.

### **1.2.3 Ansätze zur Förderung**

Erkenntnisse aus der Forschung zur diagnostischen Kompetenz sollten u.a. dazu dienen, geeignete Maßnahmen zur Förderung diagnostischer Kompetenz zu entwickeln. Mittlerweile liegen zwar erste Förderansätze vor (z.B. Förster & Souvignier, 2015; Glogger, Holzäpfel, Kappich, Schwonke, Nückles & Renkl, 2013; Klug, Gerich, Bruder & Schmitz, 2012), eine systematische Übersicht sowie umfassende Programme und entsprechende Evaluationsstudien, die sich mit der Förderung der diagnostischen Kompetenz von Lehrkräften beschäftigen, existieren bislang jedoch kaum (Artelt & Rausch, 2014).

### **1.2.4 Empirische Belege für die Relevanz diagnostischer Kompetenz**

Ein grundlegendes Problem bisheriger Forschungen zur diagnostischen Kompetenz liegt darin, dass die Bedeutung diagnostischer Kompetenz für die Qualität von Unterrichtsprozessen und die Leistungs- und Persönlichkeitsentwicklung von Schülerinnen und Schülern zwar theoretisch plausibel abgeleitet werden kann, empirisch jedoch kaum untersucht wurde (für Ausnahmen siehe Anders, Kunter, Brunner, Krauss & Baumert, 2010; Behrmann & Souvignier, 2013; Karst, Schoreit & Lipowsky, 2014; Lingelbach, 1995; Schrader, 1989). Dies ist hinsichtlich der Forderung nach messbaren Wirkungen von Unterricht und Schule (vgl. z.B. Terhart, 2002) zu kritisieren: Demnach ist es nur dann sinnvoll, diagnostische Kompetenz als wichtigen Bestandteil der Lehrerprofessionalität zu propagieren, diese zu untersuchen und in einem nächsten Schritt zu fördern, wenn die Bedeutung diagnostischer Kompetenzen für die Qualität unterrichtlicher Prozesse sowie die kognitive und motivational-emotionale Entwicklung von Schülerinnen und Schülern auch empirisch nachgewiesen ist (siehe auch Praetorius et al., 2010).

Die lediglich geringen Effekte diagnostischer Kompetenz in den bislang vorhandenen Untersuchungen könnten darauf zurückzuführen sein, dass die Leistungsentwicklung von Schülerinnen und Schülern multideterminiert ist (zsf. Hattie, 2009). Zudem spielen vermutlich vermittelnde Variablen zwischen der akkuraten Einschätzung von Merkmalen und der Lern- und Leistungsentwicklung von Schülerinnen und Schülern eine Rolle (z.B. die Umsetzung von Diagnosen in entsprechendes Lehrerhandeln; für erste Hinweise dazu siehe Behrmann & Souvignier, 2013; Helmke & Schrader, 1987), deren verstärkter Einbezug eine wichtige Aufgabe für zukünftige Forschung sein sollte. Die aktuell verbreitete Operationalisierung diagnostischer Kompetenz über die Urteilsakkuratheit lässt zudem Aspekte unberücksichtigt, die für die

praktische Unterrichtsarbeit bedeutsam sind und unter deren Einbezug größere Effekte diagnostischer Kompetenz zu erwarten wären.

### 1.3 Die Idee und Struktur des vorliegenden Bandes

Das wissenschaftliche Netzwerk zur diagnostischen Kompetenz von Lehrkräften (NeDiKo) setzte an diesen Kritikpunkten an. Ziel war es, diejenigen Desiderata zu bearbeiten, denen durch theoretische Überlegungen begegnet werden kann, um eine Grundlage für zukünftige empirische Studien zu bieten. Vier zentrale Aspekte wurden im Rahmen der Netzwerkarbeit bearbeitet: die Konzeptualisierung, die Erfassung, die Analyse und die Förderung diagnostischer Kompetenz. Im Vordergrund der Auseinandersetzungen stand also das Konstrukt diagnostische Kompetenz als solches. Auf den Bezug zum unterrichtlichen Handeln wird an vielen Stellen des Buchs eingegangen. Desiderata, die lediglich durch empirische Untersuchungen zu klären sind, blieben jedoch bei der Netzwerkarbeit ausgeklammert. Daher wird die Relevanz diagnostischer Kompetenz im vorliegenden Band nicht vertiefend behandelt.

Da die *Konzeptualisierung diagnostischer Kompetenz* grundlegend für alle weiteren Auseinandersetzungen mit dem Thema ist, lag der Schwerpunkt der Netzwerkarbeit auf diesem Bereich. Dementsprechend wird diesem Thema gegenüber den anderen Themen in diesem Band auch besonders viel Platz eingeräumt. Zunächst erfolgt eine Darstellung theoretischer Modelle, die bisher als Grundlage der Beschreibung und Erforschung diagnostischer Kompetenz dienen (Karst & Förster, Kap. 2). Das Kapitel zielt darauf ab, einen Überblick über vorhandene theoretische Ansätze zu vermitteln und zeigt die unterschiedlichen Perspektiven auf, aus denen das Thema bisher beforcht wurde. Daraufhin erfolgt eine Auseinandersetzung mit dem Kompetenzbegriff (Ufer & Leutner, Kap. 3). Anschließend wird der Kern der Netzwerkarbeit – das durch die gemeinsame Arbeit entstandene Modell diagnostischer Kompetenz – dargestellt (Herppich et al., Kap. 4). Dieses Modell kann dazu dienen, Diagnostik im schulischen Kontext zu beschreiben, und bietet eine entsprechende Reflexionsgrundlage. Es sollte in zukünftigen Forschungsarbeiten empirisch geprüft werden.

In Bezug auf die *Erfassung diagnostischer Kompetenz* wurden drei zentrale Herausforderungen identifiziert (Praetorius, Hetmanek, Herppich & Ufer, Kap. 5), die in zukünftiger Forschung zur diagnostischen Kompetenz berücksichtigt werden sollten. Zu jeder dieser Herausforderungen erfolgen in eigenen Kapiteln vertiefte Auseinandersetzungen. Karst, Klug und Ufer (Kap. 5.1) beleuchten die Notwendigkeit der Repräsentativität der ausgewählten diagnostisch relevanten Situationen und machen einen Vorschlag zu deren

Systematisierung. Mit der engen Verzahnung zwischen diagnostischem und pädagogischem Handeln setzen sich Kaiser, Praetorius, Südkamp und Ufer (Kap. 5.2) auseinander. Brühwiler (Kap. 5.3) zeigt anhand eines empirischen Beispiels auf, wie die Trennung dieser beiden Bereiche gelingen kann. Behrmann und Glogger-Frey (Kap. 5.4) geben einen Überblick über prozess- und produktbezogene Indikatoren, die das Ausmaß an diagnostischer Kompetenz erfassen und über die Feststellung der Urteilsakkuratheit hinausgehen. Eine empirische Studie, in die sowohl prozess- als auch produktbezogene Qualitätsindikatoren Eingang gefunden haben, stellen Altmann und Nückles (Kap. 5.5) vor.

In Bezug auf die *Analyse von Daten zur diagnostischen Kompetenz* wird der Fokus auf alternative Methoden zu den drei Akkuratheitskomponenten im Paradigma der Urteilsakkuratheit gelegt. Kapitel 6 richtet sich an Wissenschaftlerinnen und Wissenschaftler, die mit Daten zur Urteilsakkuratheit von Lehrkräften arbeiten und diese mittels statistischer Verfahren auswerten möchten, die über die drei Akkuratheitskomponenten hinausgehen. Karst, Hartig, Kaiser und Lipowsky (Kap. 6.1) stellen lineare Mischmodelle sowie deren praktische Anwendung mit dem frei verfügbaren Statistikprogramm R zur Analyse von Daten zur Urteilsakkuratheit vor. Humberg, Förster, Kaiser und Schönbrodt (Kap. 6.2) erläutern die Response-Surface-Analysen als eine weitere Möglichkeit, Daten zur Urteilsakkuratheit zu analysieren; auch hier erfolgt eine praktische Anwendung mit R.

In Bezug auf die *Förderung diagnostischer Kompetenz* werden verschiedene Ansätze vorgestellt, vom Eins-zu-eins-Tutoring (Herppich, Altmann, Wittwer & Nückles, Kap. 7.1), über simulationsbasierte Förderansätze (Hetmanek & van Gog, Kap. 7.2), instruktionale Ansätze (Glogger-Frey & Renkl, Kap. 7.3), den Einsatz von Tagebüchern (Klug, Kap. 7.4) bis hin zu Informationen über Lernverläufe (Förster & Souvignier, Kap. 7.5). Karing und Seidel (Kap. 7.6) diskutieren alle Beiträge im Überblick.

Ein zusammenfassender Beitrag mit Ausblick auf zukünftige Forschung (Schrader, Kap. 8) schließt den Band ab.

## 2 Ansätze zur Modellierung diagnostischer Kompetenz

*Karina Karst & Natalie Förster*

Seit langer Zeit beschäftigen sich Wissenschaftlerinnen und Wissenschaftler verschiedener Fachrichtungen mit der Einschätzung von Personen. In der Pädagogischen Psychologie und der Erziehungswissenschaft werden Einschätzungen und Urteilsprozesse von Lehrkräften über Schülerinnen und Schüler dabei häufig unter dem Oberbegriff der diagnostischen Kompetenz von Lehrkräften behandelt. Hierbei sind vielfältige Modelle entstanden, die in Publikationen zur diagnostischen Kompetenz als Grundlage verwendet oder in einschlägigen Lehrbüchern dargestellt werden. Ziel des vorliegenden Kapitels ist es, einen Überblick über diese Modelle zu geben und darauf aufbauend Implikationen für die Entwicklung eines neuen Rahmenmodells zur diagnostischen Kompetenz abzuleiten.

In der hiesigen Forschungslandschaft können vier Herangehensweisen im Umgang mit dem Konstrukt diagnostischer Kompetenz differenziert werden.

Den größten Teil machen dabei Arbeiten aus, die diagnostische Kompetenz mit der Akkuratheit von Lehrerurteilen gleichsetzen und auf die Niveau-, Differenzierungs- und Rangkomponente rekurrieren. Hierzu zählen die grundlegenden Arbeiten von Schrader (1989, 2006) sowie eine Erweiterung von McElvany et al. (2009). Beide Ansätze werden in Kapitel 2.1 von Karst vorgestellt.

Ein zentraler Kritikpunkt an dieser Perspektive auf diagnostische Kompetenz ist jedoch die Reduzierung auf die Akkuratheit und die Vernachlässigung des (bildungswissenschaftlichen) Kompetenzbegriffs.

Modelle, die den (bildungswissenschaftlichen) Kompetenzbegriff explizit mit einschließen, finden sich wiederum zum einen bei Karst (Kap. 2.2). Hier werden ausgehend vom Kompetenzbegriff nach Klieme und Leutner (2006) unterrichtliche diagnostische Anforderungssituationen abgeleitet, um die didaktische Relevanz diagnostischer Kompetenz hervorzuheben. Zum anderen wird der Kompetenzbegriff im Rahmen eines Projektes zur Bild-Text-Integration (BiTe) aufgegriffen, in dem nichtkognitiven Bestandteilen diagnostischer Kompetenz empirisch nachgegangen wird. Eine Darstellung zentraler Modellannahmen und Ergebnisse dieses Projekts erfolgt durch Ohle, McElvany, Horz und Ullrich in Kapitel 2.3.

Ein umfassenderes Verständnis für das Konstrukt diagnostischer Kompetenz zeigt sich zudem in einer Gruppe von Modellen, in denen weitere Merkmale, die über die Akkuratheit des Urteils hinausgehen, einbezogen werden. So berücksichtigen Südkamp, Kaiser und Möller (Kap. 2.4) aus einer pädagogisch-psychologischen Perspektive Lehrer-, Schüler-, Urteils- und Test-

merkmale, die die diagnostische Kompetenz ausmachen können. Gleichzeitig treffen sie Aussagen über mögliche Wirkzusammenhänge zwischen den Faktoren. Einen eher erziehungswissenschaftlichen Ansatz vertritt das Vier-Komponenten-Modell von van Ophuysen (Behrmann & van Ophuysen, Kap. 2.5). Hier wird die Qualität des Urteils erneut an der Akkuratheit festgemacht. Zusätzlich werden jedoch Qualitätsaspekte für die Daten formuliert, die der Informationserfassung und -verarbeitung sowie der Diagnose zugrunde liegen. Eine neue Sicht auf ein umfassenderes Verständnis diagnostischer Kompetenz eröffnen Glogger-Frey und Herppich (Kap. 2.6). Sie beschreiben formatives Assessment als theoretischen Ansatz zur Erforschung diagnostischer Kompetenzen von Lehrkräften und gehen damit der Frage nach, welchen Stellenwert die Lernprozessdiagnostik im Zusammenhang mit der diagnostischen Kompetenz von Lehrkräften hat.

Abschließend lässt sich eine Gruppe von Modellen unterscheiden, die den diagnostischen Prozess in den Fokus nimmt. Hier wird das Lehrerurteil nicht als Blackbox betrachtet, sondern gezielt nach dem Mechanismus der Urteilsbildung geforscht. Ein eher breit gefasstes und in vielen Disziplinen anwendbares Modell ist hierbei das Brunswick'sche Linsenmodell (Förster & Böhmer, Kap. 2.7), welches den Prozess der Informationsverarbeitung und Bedingungen für eine akkurate Urteilsbildung für nicht direkt beobachtbare Personenmerkmale beschreibt. Weiter in die Tiefe des Urteilsprozesses gehen sozialpsychologische Modelle zur (sozialen) Informationsverarbeitung (Fiske & Neuberg, 1990). Diese Modelle machen insbesondere deutlich, dass es unterschiedliche Wege für die Lehrkraft gibt (heuristisch vs. systematisch), Informationen zu integrieren und zu einem Urteil zu gelangen. Vorgestellt werden sie in Kapitel 2.8 von Böhmer, Englich und Böhmer. Nachfolgend beschreibt Klug in Kapitel 2.9 ein aktuelles Modell des Diagnoseprozesses bei Lehrkräften aus der Perspektive der Pädagogischen Psychologie. Sie unterscheidet dabei drei Phasen (präaktional, aktional und postaktional), die zyklisch verlaufen und durch verschiedene Qualitätsindikatoren gekennzeichnet sind.

Die Integration eines Rahmenmodells und eines Prozessmodells erfolgt schließlich im Ansatz von Ingenkamp und Lissmann (2008), allerdings auf einer sehr abstrakten und nicht direkt auf das Lehrerurteil anwendbaren Ebene. Behrmann und Kaiser stellen diesen sehr weit gefassten Ansatz aus der pädagogischen Diagnostik dar (Kap. 2.10).

Das Kapitel schließt mit einer kurzen Zusammenfassung und benennt Implikationen (Förster & Karst, Kap. 2.11), die sich auf Grundlage der vorgestellten Modelle für die Konzeptualisierung eines neuen Rahmen- und Prozessmodells der diagnostischen Kompetenz von Lehrkräften ergeben.

## **2.1 Akkurate Urteile – die Ansätze von Schrader (1989) und McElvany et al. (2009)**

*Karina Karst*

In der empirischen Bildungsforschung dominieren Studien zur diagnostischen Kompetenz, die sich auf die Akkuratheit des Lehrerurteils konzentrieren. In der Folge wird diagnostische Kompetenz als die zutreffende Einschätzung des Lernenden durch die Lehrkraft hinsichtlich bestimmter Merkmale definiert (Schrader, 2006).

### **2.1.1 Das Modell nach Schrader (1989)**

Im Rahmen seiner Dissertation hat Schrader (1989) erstmals in der deutschen Bildungsforschung das Konstrukt der diagnostischen Kompetenz von Lehrkräften untersucht. Motiviert war er u.a. dadurch, dass in der pädagogischen Diagnostik unterschiedliche Urteilstendenzen von Lehrkräften im Fokus standen anstelle der Genauigkeit des Lehrerurteils. Denn von der Genauigkeit – im Sinne einer zutreffenden Informiertheit – wurde ein entscheidender positiver Zusammenhang zur Anpassung der Lernumgebung an die Lernenden vermutet (Weinert & Schrader, 1986). Diese zutreffende Informiertheit mit der Abwesenheit von Urteilsfehlern gleichzusetzen, sei ein Schluss, der nicht gezogen werden dürfe, da der Zusammenhang zwischen Urteilsgenauigkeitsmaßen und verschiedenen Urteilsfehlern sehr gering ist ( $r = .05$ ), wie Murphy und Balzer (1989) in einer Metaanalyse zeigen konnten. Daraufhin differenziert Schrader (1989) drei Arten von Einschätzungen, die sich im Urteilsgegenstand unterscheiden: personenbezogene, aufgabenbezogene und aufgabenspezifische Einschätzung. Beim personenbezogenen Urteil werden leistungsbezogene Merkmale von Schülerinnen und Schülern eingeschätzt, beim aufgabenbezogenen Urteil die Schwierigkeit einer Aufgabe. Im aufgabenspezifischen Urteil wird binär bewertet, ob die oder der Lernende eine bestimmte Aufgabe richtig oder falsch löst.

Die diagnostische Kompetenz zeigt sich laut Schrader (1989) in drei Komponenten, die er für das personenbezogene und das aufgabenbezogene Urteil ermittelt. Es sind dies die Niveau-, die Differenzierungs- und die Rangordnungskomponente, deren inhaltliche Bedeutung je nach Urteilsgegenstand variiert (siehe Tabelle 1).

Mit der Niveauebene können Aussagen darüber getroffen werden, inwiefern die Lehrkraft das mittlere Leistungsniveau der Klasse oder einzelner Lernender einschätzen kann (Optimalwert = 0). Werte kleiner null ver-

weisen auf eine Unterschätzung des Niveaus, Werte größer null auf eine Überschätzung. Die Kenntnis über die Streuung der Schülerleistung wird mit der Differenzierungskomponente ermittelt. Dabei informiert das Verhältnis zwischen tatsächlicher und von der Lehrkraft eingeschätzter Streuung über die Überschätzung oder die Unterschätzung der Streuung vonseiten der Lehrkraft. Der Optimalwert ist 1. Die Rangordnungskomponente schließlich beschreibt die Korrelation zwischen Schülerleistung und Lehrerurteil. Je höher der Zusammenhang ist, desto besser kann die Lehrkraft die leistungsbezogene Rangreihe ihrer Schülerinnen und Schüler einschätzen.

Tabelle 1: Überblick über die Komponenten diagnostischer Kompetenz nach Schrader (1989) und ihre Bedeutung.

Komponente	Art des Urteils	
	aufgabenbezogen	personenbezogen
diagnostischer Kompetenz	„Wie viele Lernende werden diese Aufgabe lösen?“	„Wie viele Aufgaben werden vom Schüler/von der Schülerin richtig gelöst?“
	Akkuratheit in der Einschätzung ...	
Niveau $\bar{y} - \bar{x}$	... der mittleren Aufgabenschwierigkeit	... der mittleren Schülerleistung
Differenzierung $\sigma_y / \sigma_x$	... der Heterogenität von Aufgabenschwierigkeiten	... der Heterogenität von Schülerleistungen
Rangordnung $r_{xy}$	... der Rangreihe der Aufgabenschwierigkeiten	... der Rangreihe der Schülerleistungen

*Anmerkungen.*  $\bar{y}$  = mittleres Lehrerurteil,  $\bar{x}$  = mittlere Schülerleistung,  $\sigma_y$  = Streuung des Lehrerurteils,  $\sigma_x$  = Streuung der Schülerleistung,  $r_{xy}$  = Produkt-Moment-Korrelation zwischen Lehrerurteil und Schülerleistung.

Aufgrund der Metrik von Niveau- und Differenzierungskomponente eignen sich beide nicht für die Überprüfung gerichteter Zusammenhänge (im Sinne von: je höher, desto besser). Deswegen führte Schrader (1989) noch den Niveaufehler und den Differenzierungsfehler ein. Die Fehler entsprechen den Absolutbeträgen der jeweiligen Differenzen zwischen den mittleren Urteilen bzw. Schülerleistungen respektive der (geschätzten und realen) Leistungsstreuungen.

Für das aufgabenspezifische Urteil wurden die drei Komponenten wegen einer zu geringen Urteilsstichprobe von Schrader selbst nicht berechnet, obgleich dies natürlich bei ausreichendem Stichprobenumfang möglich ist.

## 2.1.2 Die Erweiterung von McElvany et al. (2009)

In der erweiterten Systematisierung diagnostischer Kompetenzen von McElvany et al. (2009) werden ebenso die drei Komponenten nach Schrader (1989) unterschieden. Die Erweiterung besteht in einer systematischen Trennung von Urteilsbereichen und Urteilebenen, in denen Lehrerurteile verortet sein können. Als Urteilsbereiche werden Lernende und Unterrichtsmaterial differenziert. Daneben wird jeder Urteilsbereich in drei Urteilebenen eingeteilt, die von einer sehr spezifischen Ebene (Urteile über Schülerinnen und Schüler/Aufgaben) hin zu einer globalen Ebene (Urteile über die Klassenstufe/Lernbereiche) reichen. Je Urteilebene sind es also unterschiedliche Urteilsgegenstände, die im Fokus stehen.

Für jede Ebene und jeden Bereich können theoretisch die Niveau-, Differenzierungs- und Rangordnungskomponente ermittelt werden. Dies gelingt aber nur, wenn Urteil und Kriterium die gleiche Skala aufweisen, da dies zur Berechnung der Niveau- und Differenzierungskomponente notwendig ist.

## 2.1.3 Empirische Befunde

Inwiefern die drei Komponenten zur Erfassung der Urteilsakkuratheit geeignet sind, wurde schon oft untersucht. Kontinuierlich zeigt sich, dass die Komponenten innerhalb einer Urteilsdimension (aufgabenbezogen, aufgabenspezifisch, personenbezogen), aber auch zwischen den Urteilsdimensionen nur schwach bis gar nicht korrelieren (Karst, 2012; Schrader, 1989; Spinath, 2005; Südkamp, Möller & Pohlmann 2008). Demnach müssen wohl mehrere, voneinander unabhängige Komponenten diagnostischer Kompetenz angenommen werden.

Weiterhin wird die Frage diskutiert, ob es sich bei Niveau- und Differenzierungskomponente um Merkmale handelt, die originär der Lehrerin bzw. dem Lehrer zugeschrieben werden können. Die Tatsache, dass insbesondere die personenbezogene Niveau- und Differenzierungskomponente gleichermaßen mit dem Lehrerurteil wie mit dem Kriterium (Schülerleistung) in einer Wechselbeziehung stehen, spricht dagegen (Schrader, 1989). Nur die Rangordnungskomponente sei auf dieser Grundlage als originäres Kompetenzmerkmal der Lehrkraft charakterisierbar (Schrader, 1989). In der Folge wird in neueren Studien immer wieder die These aufgegriffen, dass die Rangordnungskomponente der eigentliche Indikator diagnostischer Kompetenz sei (Helmke, Hosenfeld & Schrader, 2004; Lorenz & Artelt, 2009; Schrader & Helmke, 1987; Schrader, 2006).

Um die Auswirkungen der Urteilsakkuratheit auf die Lernentwicklung zu untersuchen, kam demnach insbesondere die Rangordnungskomponente zum Einsatz. Die Befunde sind uneinheitlich. Während Schrader (1989) keinen direkten Einfluss der diagnostischen Kompetenz feststellen konnte, präzisieren einige Komponenten die Lernentwicklung positiv. In der Studie von Anders, Kunter, Brunner, Krauss und Baumert (2010) betrifft dies den aufgabenbezogenen Urteilsfehler und die personenbezogene Rangordnungskomponente, in der Studie von Karing, Pfof und Artelt (2011b) die aufgabenspezifische Trefferquote. Dagegen zeigt sich bei Karst, Schoreit und Lipowsky (2014) ein negativer Effekt der personenbezogenen Rangordnungskomponente auf die Lernentwicklung der Lernenden, während sich alle weiteren Komponenten nicht signifikant auf die Lernentwicklung auswirken.

Hinsichtlich der Interaktion mit Unterrichtsmerkmalen konnten Schrader und Helmke (1987) ermitteln, dass eine hohe diagnostische Kompetenz (personenbezogene Rangordnungskomponente) gepaart mit stark ausgeprägten strukturierenden Verhaltensweisen der Lehrkraft den Lernerfolg in Mathematik positiv beeinflusst. Karing et al. (2011b) wie auch Behrmann und Souvignier (2013) konnten einen Moderationseffekt nachweisen. So stellten Behrmann und Souvignier (2013) fest, dass die personenbezogene Rangordnungskomponente mit der Feedbackhäufigkeit interagiert. Letztere prädiizierte die Lernentwicklung im Bereich Leseverständnis erst dann, wenn die Lehrkraft über eine hohe diagnostische Kompetenz (personenbezogene Rangordnungskomponente) verfügte.

### **2.1.4 Stärken und Schwächen**

Die Komponenten von Schrader (1989) wurden und werden in zahlreichen Studien immer wieder aufgegriffen. Die Erweiterung von McElvany et al. (2009) in Form einer Systematisierung von Urteilsbereichen und Urteilsbenen liefert eine hilfreiche Taxonomie zur Verortung diagnostischer Urteile und macht damit den Vergleich zwischen gleichen bzw. unterschiedlichen diagnostischen Kompetenzen beispielweise in ihrer prädiktiven Kraft für die Lernentwicklung Lernender leichter.

Die auffallende Fokussierung auf die Rangkomponente ist jedoch aus zwei Gründen problematisch. Wenn erstens nur ein einziger Aspekt der Urteilsakkuratheit in Betracht gezogen wird, bleibt die Frage nach der Struktur, Generalität und Spezifität diagnostischer Kompetenz theoretisch und empirisch offen. Die Klärung dieser Frage sei aber, so Schraders (1989) Schlussfolgerung, ein wichtiges und notwendiges Unterfangen. Zudem werden zwei-

tens bei einer Urteil-Kriterium-Korrelation über mehrere Aufgaben und Lernende hinweg Mittelwerts- und Streuungsunterschiede außer Acht gelassen, was zu einer Überschätzung diagnostischer Kompetenz führen kann (Bates & Nettelbeck, 2001; Begeny, Eckert, Montarello & Storie, 2008).

Weiterhin kann die operationale Definition diagnostischer Kompetenzen kritisch hinterfragt werden. Hierdurch wird sowohl im Ansatz von Schrader (1989) als auch in der Erweiterung von McElvany et al. (2009) die didaktische Relevanz diagnostischer Kompetenz nicht berücksichtigt (Abs, 2007; Karst, 2012). Es bleibt beispielsweise offen, in welchem spezifischen Kontext (unterrichtliche Situation) solche Diagnoseleistungen erbracht werden müssen.

## 2.2 Diagnostische Kompetenz und unterrichtliche Situationen

*Karina Karst*

Wenn bislang von diagnostischer Kompetenz die Rede war, fällt auf, dass der bildungswissenschaftliche Kompetenzbegriff wenig berücksichtigt wurde. Karst (2012) unternimmt hierzu einen ersten Versuch, indem sie in ihrer Dissertation diagnostische Anforderungen und Situationen von Lehrkräften analysiert, die aus der alltäglichen Unterrichtspraxis resultieren. Welche Situationen im Unterricht lassen sich durch diagnostische Anforderungen kennzeichnen? Sind diese Anforderungen über unterschiedliche Situationen hinweg gleich? Zugleich soll mit dieser Anforderungsanalyse die didaktische Relevanz diagnostischer Kompetenz von Lehrkräften erhöht werden.

Als für jede Lehrkraft grundlegende diagnostische Anforderung definiert Karst (2012) die Einschätzung der kognitiven Lernvoraussetzungen der Schülerinnen und Schüler. Denn wir wissen, dass Lernprozesse besonders dann positiv verlaufen, wenn sie am Vorwissen des Lernenden ansetzen. Diese Lernprozesse finden u.a. im Unterricht statt und können damit von der Lehrkraft beispielsweise durch eine adäquate kognitive Aktivierung der Lernenden positiv beeinflusst werden (Lipowsky, Rakoczy, Pauli, Drollinger-Vetter, Kliehme & Reusser, 2009). Ausgehend vom unterrichtlichen Alltag in der Grundschule und unter Einbezug der Vorarbeiten von Schrader (1989) und McElvany et al. (2009) beschreibt Karst (2012; vgl. auch Karst et al., 2014) drei idealtypische unterrichtliche Situationen, in denen diese diagnostische Anforderung unterschiedlich bewältigt werden kann. Die Unterschiedlichkeit der Situationen ergibt sich dabei aus der Perspektive, die von der Lehrkraft bei der Einschätzung der Lernenden eingenommen werden sollte sowie vermutlich aus spezifischen Arten diagnostischen Wissens der Lehrkraft (Wissen

über Aufgaben, Wissen über Lernende; kognitive Leistungsdispositionen im Sinne der Kompetenzdiagnostik nach Klieme und Leutner (2006)), welches zur Anforderungsbewältigung aktiviert werden muss. Im Folgenden werden diese Situationen erläutert.

### *Klassenbezogene Situation*

Die klassenbezogene Situation beschreibt die unterrichtliche Phase, bei der die Lehrkraft Entscheidungen treffen muss, wie sie beispielsweise in ein neues Thema einführen oder die Klasse auf Unterrichtsinhalte einstimmen kann (Kann das Thema schon aufgegriffen werden? Ist ausreichend Vorwissen vorhanden?). Dabei ist die Klasse als Ganzes im Fokus, da die Lehrkraft das Vorwissen aller Lernenden berücksichtigen sollte. Es geht um die Auswahl von Aufgaben, die für alle Schülerinnen und Schüler geeignet sind. Das diagnostische Urteil wird über die Merkmalsträger Klasse und Aufgabe bzw. Inhaltsbereich gefällt.

Zur erfolgreichen Bewältigung der geplanten Aufgaben benötigt die Lehrkraft Wissen über das allgemeine Leistungsniveau der Lernenden (Niveauwissen). Dies bezeichnet Karst (2012) als niveaubezogene diagnostische Kompetenz. Daneben muss die Lehrkraft mögliche Alternativaufgaben stellen können. Ihr muss die relative Position der Aufgabenschwierigkeit gegenüber anderen Aufgaben aus diesem Inhaltsbereich bewusst sein. Dies kommt in der aufgabenheterogenitätsbezogenen diagnostischen Kompetenz zum Ausdruck.

### *Schülerglobale Situation*

In dieser modellhaft formulierten Situation stehen einzelne Schülerinnen und Schüler im Fokus des Lehrerurteils. Eine solche diagnostische Situation ist denkbar, wenn die Lehrkraft in einer Übungs- oder Vertiefungsphase auf bindend differenzierende Maßnahmen zurückgreift. Dabei wird zumeist leistungsbezogen differenziert auf Grundlage des Wissens der Lehrkraft über die relative Position der Lernenden (schülerheterogenitätsbezogene diagnostische Kompetenz). Den Gruppen, die sich im Leistungsniveau unterscheiden (bei leistungshomogener Gruppierung), können sodann verschiedene Aufgaben zugeteilt werden, die in der Quantität und/oder im Schwierigkeitsgrad voneinander abweichen (aufgabenheterogenitätsbezogene diagnostische Kompetenz). Die Annahme ist hier, dass die Lehrkraft mit dem Ziel der Gruppenbildung eine globale Sicht auf den Lernenden einnimmt. Gleichzeitig ist aber auch die fachdidaktische Einschätzung der einzusetzenden Aufgaben notwendig. Die Merkmalsträger, über welche geurteilt wird, sind der Lernende im Allgemeinen und die Aufgabe/der Inhaltsbereich.

### *Schülerspezifische Situation*

In dieser Situation geht es um spezifische Urteile bei einzelnen Schülerinnen und Schülern. Dieser Fall tritt ein, wenn die Lehrkraft Lernende individuell fördern möchte, beispielsweise durch die Zuweisung geeigneter Aufgaben, Hilfsangebote, durch zusätzliche individuelle Erläuterungen oder durch den gezielten Einsatz von Veranschaulichungsmitteln. Dabei muss die Lehrkraft über einzelne Stärken oder Schwächen eines Lernenden informiert sein (schülerprofilbezogene diagnostische Kompetenz). Hier wird ein schülerspezifisches Wissen aktiviert, welches durch die Interaktion zwischen Lernendem und Aufgabe entsteht. So weiß die Lehrkraft beispielsweise, dass Schüler A die kognitiv anspruchsvolle Aufgabe X ohne Weiteres bearbeiten kann, während Schülerin B die Aufgabe X schwerfällt. Schülerin B kann dafür Aufgabe Y gut bearbeiten, was Schüler A nicht gelingt. Im Gegensatz zur schülerglobalen Diagnosesituation fokussiert die Lehrkraft hier also auf die Lernvoraussetzungen Einzelner bei bestimmten Lerninhalten.

Durch die so – idealtypisch – formulierten Situationen wird der Forderung einer didaktischen Einbettung der Beurteilungssituation Rechnung getragen.

## **2.2.1 Empirische Befunde**

Das theoretisch formulierte Modell von Karst (2012) wurde bislang insbesondere mit Blick auf drei Fragestellungen empirisch analysiert.

*Bilden die situationsspezifischen Teilkompetenzen in ihrer Gesamtheit ein übergeordnetes Konstrukt diagnostischer Kompetenz ab?*

In der Dissertation von Karst (2012) zeigen sich keine bedeutsamen Zusammenhänge, die einen Hinweis auf ein existierendes globales Konstrukt diagnostischer Kompetenz geben (Karst, 2012, S. 192). Einzig die aufgabenheterogenitätsbezogene diagnostische Kompetenz der klassenbezogenen Diagnosesituation korreliert signifikant positiv ( $r = .34$ ,  $p < .05$ ) mit der gleichen Teilkompetenz der schülerglobalen Diagnosesituation.

*Können auf Grundlage schülerspezifischer Urteile Aussagen über die schülerglobale und/oder klassenbezogene diagnostische Kompetenz gemacht werden?*

Zur Überprüfung dieser Frage wurden spezifische Einzelurteile auf eine höhere Ebene aggregiert, sodass beispielsweise ein schülerglobales Urteil resultiert (Karst, 2012). Geprüft wurden die Korrelationen von gleichen origina-

len Teilkompetenzen mit solchen Teilkompetenzen, die anhand aggregierter Urteile ermittelt wurden. Es zeigen sich keine signifikanten Zusammenhänge (Karst, 2012, S. 209), sodass auf Grundlage schülerspezifischer Urteile keine Aussagen über die schülerglobale und klassenbezogene diagnostische Kompetenzen gemacht werden können.

*Wie prädiktiv sind die situationsspezifischen diagnostischen Kompetenzen für die Lernentwicklung von Grundschulern im Fach Mathematik?*

Im Artikel von Karst et al. (2014) zeigen sich differenzielle Effekte der situationsspezifischen diagnostischen Kompetenzen. Während die schülerspezifische Diagnosekompetenz die Lernentwicklung der Schülerinnen und Schüler nicht prädiziert, ergibt sich für die klassenbezogene Diagnosekompetenz der Lehrkraft ein signifikant positiver Effekt. Erwartungswidrig sagt die schülerglobale Diagnosekompetenz die Lernentwicklung signifikant negativ vorher.

Die Befunde verweisen auf die Notwendigkeit einer nach Situationen getrennten und spezifischen Erfassung diagnostischer Urteile und diagnostischer Kompetenzen.

### **2.2.2 Stärken und Schwächen**

Die Herangehensweise von Karst (2012) bringt den Vorteil mit sich, dass durch die Anerkennung des kontextuellen Charakters von Kompetenzen ein zentrales Bestimmungsstück des Kompetenzbegriffs nach Klieme und Leutner (2006) berücksichtigt wird, auf welchen sich auch die Arbeiten des wissenschaftlichen Netzwerks zur diagnostischen Kompetenz von Lehrkräften (NeDiKo) beziehen.

Eine Unterscheidung von Situationen im pädagogischen Handlungsfeld wäre jedoch nicht notwendig, wenn angenommen werden könnte, dass die diagnostische Kompetenz der Lehrkraft über mehrere Situationen hinweg dieselbe ist. Davon ist aber nicht auszugehen. Es gibt Hinweise darauf, dass sich die Diagnosekompetenz je nach Unterrichtsfach und einzuschätzendem Merkmal unterscheidet (Coladarci, 1986; Praetorius, Karst, Dickhäuser & Lipowsky, 2011; Spinath, 2005). Auch die bei Karst (2012) geringen bis nicht vorhandenen Korrelationen zwischen gleichen diagnostischen Teilkompetenzen aus unterschiedlichen Situationen deuten darauf hin.

Kritisch muss jedoch die idealtypische Herleitung der diagnostischen Situationen betrachtet werden. Was spricht für die Diagnosesituationen im Unterricht? Handeln Lehrkräfte tatsächlich so? Sind weitere Situationen denkbar? Sicherlich handelt es sich bei den Diagnosesituationen um relativ plausible

und verallgemeinerbare Unterscheidungsmöglichkeiten. Dennoch ist eine Validierung im unterrichtlichen Kontext notwendig, beispielsweise durch die Analyse von Unterrichtsvideos oder Experteninterviews. Auch sind weitere Diagnosesituationen zu berücksichtigen, die prozessbezogene Urteile im Sinne formativen Assessments enthalten oder mehr im Zusammenhang mit formellen Diagnosen (z.B. Noten, Übergangsentscheidungen) stehen. Wie eine solche Systematisierung von Situationen aussehen kann, wird bei Karst, Klug und Ufer (Kap. 5.1) dargestellt.

### **2.3 Aspekte diagnostischer Kompetenz im Projekt „BiTe“**

*Annika Ohle, Nele McElvany, Holger Horz & Mark Ullrich*

Dieses Subkapitel nimmt professionelle Merkmale von Lehrenden als Voraussetzung für diagnostische Handlungen in den Blick und fokussiert dabei auf Aspekte von Lehrkraftkompetenz, die spezifisch auf unterrichtsbezogene Diagnostik übertragen wurden. Im Rahmen des DFG-geförderten Projekts „Entwicklung und Überprüfung von Kompetenzmodellen zur integrativen Verarbeitung von Texten und Bildern“ (kurz: Bild-Text-Integration; BiTe; Ohle et al., in Druck) wurden Aspekte diagnostischer Lehrkraftkompetenz beim Unterrichten mit Bild-Text-Material erfasst. Neben kognitiven Dispositionen wurden hierbei auch affektive, motivationale und selbstbezogene Kompetenzen als Voraussetzungen für diagnostische Urteile berücksichtigt (Weinert, 2001). Die diagnostische Kompetenz ist in diesem speziellen Bereich von besonderer Relevanz, da die Integration von Informationen aus zwei unterschiedlichen Quellen (Bild und Text) eine zusätzliche kognitive Herausforderung für Lernende darstellt (Brünken, Plass & Leutner, 2003). Die diagnostische Kompetenz von Lehrkräften ist ein zentraler Aspekt ihrer Professionalität und wird als wichtige Voraussetzung angenommen, um adäquate Lernangebote für Schülerinnen und Schüler zu gestalten (Ohle & McElvany, 2015).

Neben den kognitiven Leistungsdispositionen, wie sie auch im Kompetenzansatz von Leutner und Klieme (2006) postuliert werden, sind für das kompetente Handeln einer Lehrkraft ebenso nichtkognitive, kontextspezifische Handlungsdispositionen erforderlich. Ihre Bedeutung konnte bereits in verschiedenen empirischen Untersuchungen bestätigt werden (Baumert & Kunter, 2006). An dieser Stelle ist die Studie „Professionswissen von Lehrkräften, kognitiv aktivierender Mathematikunterricht und die Entwicklung mathematischer Kompetenz“ (COACTIV) zu nennen. Im COACTIV-Modell zur professionellen Handlungskompetenz von Lehrkräften sind neben dem