



Diagnose von Bewertungskompetenz durch Biologielehrkräfte – Negieren eigener Fähigkeiten oder Bewältigen einer Herausforderung?

Benjamin Steffen · Corinna Höble

Eingegangen: 12. Dezember 2014 / Angenommen: 13. Juli 2015 / Online publiziert: 5. September 2015

© Gesellschaft für Didaktik der Physik und Chemie (GDPC); Fachsektion Didaktik der Biologie im VBIO (FDdB im VBIO) and Springer-Verlag GmbH Berlin Heidelberg 2015

Zusammenfassung Die Integration des durch die Bildungsstandards vorgegebenen Kompetenzbereichs ‚Bewertung‘ in den Biologieunterricht ist mit Schwierigkeiten für Lehrende und Lernende verbunden. Insbesondere das Diagnostizieren von Schülerleistungen im Bereich ethischen Bewertens durch Lehrkräfte stellt in diesem Kontext einen Problembereich dar. Davon ausgehend befasst sich die vorliegende Studie mit Deutungs- und Handlungsmustern von Biologielehrkräften zur Diagnose von Bewertungskompetenz. Die Datenerhebungen im Rahmen dieser qualitativen Studie beinhalten die von zwei Interviews mit der jeweiligen Lehrkraft eingerahmte Videographie einer Unterrichtsstunde zur Förderung von Bewertungskompetenz. Die Auswertung der Interviewdaten erfolgte nach den Verfahren der Grounded Theory. Die zentrale Kategorie im Hinblick auf die Diagnose von Bewertungskompetenz wurde als „Negiertes Bewältigen“ konzeptualisiert. Dieses Konzept beschreibt die Erscheinung, dass die interviewten Lehrkräfte Unsicherheiten und Schwierigkeiten bei der Diagnose von Bewertungskompetenz nannten oder sich ihre diesbezüglichen Fähigkeiten ganz absprachen, gleichzeitig jedoch in der Lage waren, wesentliche Aspekte von Bewertungskompetenz in ausgewählten Videovignetten zu diagnostizieren. Das Phänomen wird hinsichtlich seiner ursächlichen, kontextuellen und intervenierenden Bedingungen, der Handlungsstrategien und Konsequenzen beschrei-

bend erklärt. Aus den Ergebnissen werden Implikationen für die Lehrerbildung abgeleitet.

Schlüsselwörter Diagnose · Bewertungskompetenz · Socioscientific issues · Qualitative Forschung · Grounded Theory

Assessment of decision-making in socioscientific issues by teachers: Negating one’s own abilities or mastering a challenge?

Abstract The integration of socioscientific issues into biology lessons constitutes a real challenge for both, biology teachers and learners. Especially the assessment of students’ performances regarding socioscientific issues provides difficulties. Against this background this study’s research question focuses on biology teachers’ interpretive patterns and action patterns regarding the assessment of students’ performances in socioscientific issues. During data acquisition interviews with biology teachers were conducted, followed by a videotaped lesson about a socioscientific issue in a class of the interviewed teacher, and a second interview afterwards. Analysis of the interview data followed the strategies of grounded theory. The core category regarding the assessment of students’ performances in socioscientific issues was named “negated mastering”. This concept describes the phenomenon, that teachers stated insecurities and difficulties in the context of assessment of socioscientific issues or even indicated a lack of own abilities, but, nevertheless, were able to identify and assess main aspects of students’ performances in socioscientific issues. This concept is described regarding its causal, contextual and intervening conditions, its (inter-) actions and

B. Steffen (✉) · C. Höble
Institut für Biologie und Umweltwissenschaften, Arbeitsgruppe
Didaktik der Biologie, Carl von Ossietzky Universität Oldenburg,
Carl-von-Ossietzky-Straße 9–11,
26111 Oldenburg, Deutschland
E-Mail: benjamin.steffen@uni-oldenburg.de

consequences. Based on the results, guidelines for teacher education are developed.

Keywords Assessment · Decision-making · Socioscientific issues · Qualitative research · Grounded theory

Einleitung

Gentechnisch veränderte Lebensmittel, Präimplantationsdiagnostik, alternative Formen der Energiegewinnung oder Eingriffe des Menschen in Ökosysteme bilden Themen, die in unterschiedlichem Ausmaß Bezüge zum Bereich der Biologie und gleichzeitig eine hohe gesellschaftliche Relevanz besitzen, da sie Wertesysteme moderner Gesellschaften tangieren. Die moderne Biologie beinhaltet somit verschiedene Themengebiete, die Verbindungen der Wissenschaft mit gesellschaftlichen Ebenen aufweisen. Folglich kann sich ein moderner Biologieunterricht nicht ausschließlich der Vermittlung fachlichen Wissens widmen (Bögeholz et al. 2004). Diesen Gedanken greifen die von der Kultusministerkonferenz (KMK 2005) verabschiedeten Bildungsstandards auf, indem sie für das Fach Biologie mit den Domänen ‚Fachwissen‘, ‚Erkenntnisgewinnung‘, ‚Kommunikation‘ und ‚Bewertung‘ vier übergeordnete Kompetenzbereiche vorgeben, in denen Schüler Fähigkeiten und Fertigkeiten erwerben sollen.

Die Integration von Thematiken, durch welche sich Kompetenzen der Bewertung erwerben und fördern lassen, beinhaltet für Lehrkräfte der Biologie größere Schwierigkeiten (Lewis 2006; Alfs 2012). Insbesondere das Diagnostizieren von schülerseitigen Bewertungsprozessen bereitet Lehrkräften der Biologie oftmals Probleme (Sadler 2011; Alfs 2012; Fensham und Rennie 2013). Gezielt empirisch untersucht wurden diese Probleme jedoch bislang kaum. Vor diesem Hintergrund will die vorliegende Untersuchung aus der Perspektive von Lehrkräften auf die Diagnose von Schülerleistungen zum Kompetenzbereich Bewertung fokussieren. Auf der Basis problemzentrierter Interviews (Witzel 2000) und aus videographierten Unterrichtsstunden zum Themenkomplex Klimawandel ausgewählten Vignetten sollen mittels der Kodiervorgang der Grounded Theory (Strauss und Corbin 1996) Deutungs- und Handlungsmuster von Lehrkräften zu einer Theorieskizze zur Diagnose von Bewertungskompetenz verdichtet werden. Die globale Erwärmung bildet als aktuelles und alltagsrelevantes Thema den Bewertungskontext dieser Studie. Aus der entwickelten Skizze – dem als „Grounded Theory“ bezeichneten Ergebnis einer gleichnamigen Studie – werden Implikationen für die Aus- und Weiterbildung von Lehrkräften hinsichtlich der Diagnose von Bewertungskompetenz abgeleitet.

Theoretische Grundlagen

Fachdidaktisches Wissen als Teil des Professionswissens

Das Professionswissen bildet den Kern der beruflichen Kompetenz von Lehrpersonen und stellt im Diskurs über Qualitätssicherung und -verbesserung der deutschen Bildungslandschaft einen aktuellen und hoch relevanten Gegenstand dar (Artelt und Gräsel 2009; Tepner et al. 2012). Es gliedert sich als ein komplexes Konstrukt in unterschiedliche Bereiche und schärft sich beginnend während der Ausbildung, weiter ausdifferenziert dann aber vor allem durch praktische Berufserfahrung aus (van Driel et al. 1998). In Bezug auf die Topologie des Professionswissens findet heute eine Unterteilung in die drei Domänen pädagogisches Wissen (pedagogical knowledge), Fachwissen (content knowledge) und fachdidaktisches Wissen (pedagogical content knowledge) im Anschluss an Shulman (1987) verbreitet Verwendung. Alle drei Bereiche bilden wesentliche Komponenten bei der erfolgreichen Gestaltung von Lernprozessen (Blömeke et al. 2009); gleichzeitig finden sie ihre Entsprechungen in den pädagogischen, fachlichen resp. fachdidaktischen Bereichen der Lehrerbildung in Deutschland (Blömeke et al. 2008). Im Lehrberuf bildet die Komponente des fachdidaktischen Wissens als Bestandteil des Professionswissens somit ein zentrales Element. Fachdidaktisches Wissen ist notwendig, um fachspezifische Inhalte so aufzubereiten und zu organisieren, dass Lernende sich diese aneignen können. In der Literatur wird das fachdidaktische Wissen als Wissensfacette des Professionswissens von Lehrkräften überwiegend angelehnt an die Definition von Shulman (1987, S. 8) beschrieben. Er charakterisiert das Konstrukt als „the blending of content and pedagogy into an understanding of how particular topics, problems, or issues are organized, represented, and adapted to the diverse interests and abilities of learners, and presented for instruction.“ Hinsichtlich der Unterteilung des fachdidaktischen Wissens in einzelne Wissensfacetten herrscht in der Literatur kein Konsens, auch wenn sich postulierte Unterteilungen oftmals nur in einzelnen Details unterscheiden. So beschreiben Park und Oliver (2008) im Rekurs auf van Driel et al. (1998) fünf Wissensfacetten des fachdidaktischen Wissens: ‚Orientations to teaching science‘, ‚knowledge of students‘ understanding in science‘, ‚knowledge of science curriculum‘, ‚knowledge of instructional strategies and representations for teaching science‘ sowie ‚knowledge of assessment of science learning‘. Demnach wird die Diagnose von Lernprozessen in der letztgenannten Facette als Teil des fachdidaktischen Wissens eingeordnet. „Fachübergreifende Prinzipien des Diagnostizierens, Prüfens und Bewertens“ (Baumert und Kunter 2006, S. 485) werden dagegen dem allgemeinen pädagogischen Wissen zugeordnet.

Diagnostizieren als Teil des fachdidaktischen Wissens

Diesen im fachdidaktischen Wissen verorteten diagnostischen Fähigkeiten von Lehrkräften wird „in der bildungspolitischen Diskussion der letzten Jahre in der Bundesrepublik die Rolle einer wichtigen Stellschraube der Qualitätsentwicklung zugeordnet“ (Artelt und Gräsel 2009, S. 157). In vielen Studien ist dabei der Begriff „diagnostische Kompetenz“ zentral, ohne dass sich auf ein entsprechendes theorie- oder empiriebasiertes Kompetenzmodell bezogen wird. Entsprechend wird der Begriff der diagnostischen Kompetenz z. T. als „irreführend“ bezeichnet, da „mit ihm eine übergreifende Fähigkeit, sichere Urteile zu fällen, suggeriert wird“ (Karst 2012, S. 215).

Verschiedene Autoren unterscheiden hinsichtlich des unterrichtlichen Diagnostizierens neben der Differenzierung zwischen Selektions- und Förderdiagnostik (Abs 2006) zwischen Formen von formellen (auch: expliziten oder wissenschaftlichen) und informellen (auch: impliziten oder alltagsnahen) diagnostischen Prozessen (Helmke 2009; Hesse und Latzko 2011; Hesse 2014). Formelle Diagnosen entstehen bewusst, systematisch und orientieren sich an wissenschaftlichen Gütekriterien. Informelle oder implizite Diagnosen werden dagegen „eher beiläufig und unsystematisch im Rahmen des alltäglichen erzieherischen Handelns“ (Helmke 2009, S. 122) erlangt. Hesse (2014) betont dabei, dass in vielen Bereichen des unterrichtlichen Alltags eine implizite Diagnostik ausreichend ist, solange sich die Lehrkraft der möglichen Unzulänglichkeit ihrer Diagnosen bewusst ist. Dennoch gelte es, fehleranfällige implizite Diagnosen immer wieder explizit zu überprüfen.

Sind die Ausführungen zu diagnostischen Fähigkeiten in der Literatur verhältnismäßig heterogen, so besteht bei der Betrachtung des Ergebnisses eines diagnostischen Prozesses – der Diagnose – weitgehender Konsens. Diagnosen manifestieren sich nach Hesse und Latzko (2011, S. 25) als „explizite Aussagen über Zustände, Prozesse oder Merkmale von Personen, die in einem reflektierten und methodisch kontrollierten Prozess gewonnen werden“ und sollen im Rahmen der vorliegenden Untersuchung als solche verstanden werden.

Für das schulische Lernen und den Unterricht besitzen die diagnostischen Fähigkeiten der Lehrkraft insofern eine große Bedeutung, als dass „die Schwierigkeit von Unterrichtsmaßnahmen, Fragen und Aufgaben auf die Lernvoraussetzungen der Schülerinnen und Schüler abgestimmt sein muss“ (Helmke 2009, S. 123–124). Die Lernvoraussetzungen der Schüler in den Blick nehmend, sollen diagnostische Fähigkeiten der Lehrkraft dazu dienen, für möglichst viele Schüler optimale Ergebnisse hinsichtlich des Erreichens von Unterrichtszielen zu ermöglichen. Dies gilt ebenso für den Erwerb von Kompetenzen der Bewertung, wie sie als

Ziel vieler naturwissenschaftlicher Curricula ausgewiesen werden (Jones et al. 2010).

Der Kompetenzbereich ‚Bewertung‘ im Biologieunterricht und anderen naturwissenschaftlichen Fächern: Grundlagen, Ziele und Perspektiven von Lehrkräften bezüglich Implementation und Diagnose

Der Erwerb von Kompetenzen im naturwissenschaftlichen Unterricht soll eine naturwissenschaftliche Grundbildung auf Seiten der Lernenden ausbilden. Diese soll „dem Individuum eine aktive Teilhabe an gesellschaftlicher Kommunikation und Meinungsbildung über technische Entwicklung und naturwissenschaftliche Forschung“ (KMK 2005, S. 6) ermöglichen. Kompetenzen werden im Rahmen der Lehr- und Lernforschung zumeist in Bezug auf Weinert (2001, S. 27–28) beschrieben als „die bei Individuen verfügbaren oder durch sie erlernbaren kognitiven Fähigkeiten und Fertigkeiten, um bestimmte Probleme zu lösen, sowie die damit verbundenen motivationalen, volitionalen und sozialen Bereitschaften und Fähigkeiten, um die Problemlösungen in variablen Situationen erfolgreich und verantwortungsvoll nutzen zu können.“ Die Bildungsstandards geben für den Erwerb einer solchen Grundbildung im Unterricht für die naturwissenschaftlichen Fächer Bewertungskompetenz als einen der vier o. g. Kompetenzbereiche vor. Die somit verankerte Integration von Bewertungsprozessen in den naturwissenschaftlichen Unterricht ist als ein internationaler Trend zu bezeichnen. Im englischsprachigen Raum wird zumeist unter dem Stichwort „socioscientific issues“ (SSI) auf entsprechende Ansätze rekurriert (Sadler 2011).

Bezogen auf das Fach Biologie fokussiert der Kompetenzbereich ‚Bewertung‘ auf das Erkennen und Bewerten biologischer Sachverhalte in verschiedenen Kontexten (KMK 2005). Zentrales Ziel von Bewertungskompetenz ist die Fähigkeit, sich „am gesellschaftlichen, z. T. kontrovers geführten Diskurs beteiligen“ zu können (KMK 2005, S. 12). Die Beurteilung von Sachverhalten oder Handlungsoptionen soll dabei explizit unter Bezugnahme auf ethische Werte vollzogen werden. Hier wird in Bezug auf den Kompetenzbereich ‚Bewertung‘ der Standards für den Biologieunterricht die Nähe zum Bereich der Ethik in einer Form deutlich, die in den Fächern Chemie und Physik schwächer oder gar nicht ausgeprägt ist (Dittmer und Gebhard 2012). Hinweise auf eine Fachspezifität – und damit indirekt eine Kontextabhängigkeit – von Bewertungskompetenz lieferten auch Hostenbach und Walpuski (2013).

Im Bereich der biologiedidaktischen Forschung sind drei Modelle explizit für den Kompetenzbereich ‚Bewertung‘ entwickelt worden. Das Modell zur ethischen Urteilskompetenz (Reitschert et al. 2007), das Göttinger Modell der Bewertungskompetenz (Bögeholz 2007) sowie das Kompetenzmodell aus dem Projekt ‚Evaluation der Standards in

den Naturwissenschaften für die Sekundarstufe I¹ (ESNaS; Hostenbach et al. 2011) entstammen unterschiedlichen thematischen Kontexten und beziehen sich auf unterschiedliche Grundlagen (Hostenbach et al. 2011; Steffen und Höhle 2014). Das Modell der ethischen Urteilskompetenz ist ein Strukturmodell, welches in bio- bzw. medizinethischen Kontexten entwickelt wurde und den Kompetenzbereich ‚Bewertung‘ auf der Basis von bereits bestehenden Modellen zur Urteilsbildung aus Philosophie- und Biologiedidaktik beschreibt (Reitschert et al. 2007). Das Göttinger Modell der Bewertungskompetenz fokussiert thematisch auf den Bereich der nachhaltigen Entwicklung und rückt vor allem Theorien der Entscheidungsfindung in den Fokus (Bögeholz 2007). Im Rahmen dieser Studie wurden die Teilkompetenzen des Modells zur ethischen Urteilsbildung (‚Wahrnehmen und Bewusstmachen der eigenen Einstellung‘, ‚Wahrnehmen und Bewusstmachen moralischer Relevanz‘, ‚Beurteilen‘, ‚Folgenreflexion‘, ‚Urteilen‘, ‚Ethisches Basiswissen‘ und ‚Argumentieren‘) sowie Teilkompetenzen des Göttinger Modells (‚Kennen und Verstehen von Nachhaltiger Entwicklung‘, ‚Kennen und Verstehen von Werten und Normen‘, ‚Generieren und Reflektieren von Sachinformationen‘ sowie ‚Bewerten, Entscheiden und Reflektieren‘) einerseits zugrunde gelegt, um in den Videographien getätigte Schüleräußerungen zum Kompetenzbereich ‚Bewertung‘ zu identifizieren. Andererseits wurden die diagnostischen Äußerungen der Lehrkräfte zu den jeweiligen Vignetten mit den Teilkompetenzen der Modelle abgeglichen, um deren Qualität einschätzen zu können.

Einstellungen und Perspektiven von Lehrkräften hinsichtlich der Implementation von Bewertungskompetenz (bzw. SSI) in den naturwissenschaftlichen Unterricht waren bislang Gegenstand vergleichsweise weniger Studien (Dittmer 2010; Mrochen und Höttecke 2012; Hartmann-Mrochen 2013). Dass die Zuschreibung von Relevanz zu Ethik als Teil des naturwissenschaftlichen Unterrichts innerhalb der Personengruppe der Lehrer variiert, stellten Sadler et al. (2006) in einer US-amerikanischen Studie fest. Sie ermittelten innerhalb eines Samples von 22 interviewten Lehrkräften hinsichtlich der Integration von SSI in den naturwissenschaftlichen Unterricht fünf unterschiedliche Profile. Nahezu alle Lehrkräfte wiesen dabei den Bereichen von Ethik und Werten eine vergleichsweise hohe Relevanz zu, die jedoch unterschiedliche Ausprägungen annahm. Einige Lehrkräfte sahen kontroverse ethische Themen ebenso wie gesellschaftlich verhandelbare Werte als einen wesentlichen und von ihnen auch in der Praxis umgesetzten Teil des Unterrichts an. Andere stimmten mit dieser grund-

sätzlichen Einstellung überein, sahen in der Praxis jedoch viele Hindernisse, die der eigentlichen Umsetzung im Wege standen. Entsprechend verwendeten diese Lehrkräfte nur sehr wenig Zeit auf die Diskussion von SSI im Unterricht. Als primäre Gründe wurden hier Zeitmangel und normative Vorgaben genannt. Eine weitere Gruppe von Lehrkräften wies ethischen Dilemmata zwar generell eine hohe Relevanz zu, wertete diese aber nicht als prioritäre Ziele von naturwissenschaftlicher Grundbildung. Lediglich eine der 22 interviewten Lehrkräfte stritt die Relevanz von Ethik als Teil naturwissenschaftlichen Unterrichts ab. Ein letzter Typ von Lehrkräften ordnete dagegen ethische Werte vielmehr als Teil von Bildung im Allgemeinen und nicht speziell als Element im Rahmen von SSI im naturwissenschaftlichen Unterricht ein (Sadler et al. 2006).

Das Beurteilen bzw. Diagnostizieren von Bewertungskompetenz durch Biologielehrkräfte wurde bislang kaum *explizit* untersucht (Fensham und Rennie 2013; Sadler 2011). In der Studie von Ekborg et al. (2013) implementierten schwedische Lehrkräfte SSI in ihren Unterricht, behandelten innerhalb der eigenständig geplanten und durchgeführten Einheiten jedoch kaum vertieft ethische Problematiken. Die Autoren schlussfolgern, dass sich die entsprechenden Unterrichtssequenzen (obwohl als SSI indendiert) durch die Ausklammerung konstituierender Merkmale eher im Bereich der Verbindung von Wissenschaft und Gesellschaft, als in den der ethisch-moralischen SSI verorten ließen. Die Leistungen der Lernenden wurden seitens der Lehrkräfte – entgegen ihrem Vorgehen im „regulären“ Unterricht – nicht bewertet. Ratcliffe und Grace (2003) führen die Schwierigkeit der Diagnose („assessment“) im Rahmen von SSI auch auf deren Facettenreichtum zurück. Alfs (2012) stellte in einer Interviewstudie an deutschen Biologielehrkräften fest, dass die Relevanz des Kompetenzbereichs für den Biologieunterricht – und damit auch die Bedeutung desselben für die Notengebung – von den interviewten Lehrkräften durchaus wahrgenommen, aber nach eigenen Auskünften in der Praxis zumeist nicht umgesetzt wird. Die von den Lehrkräften beschriebenen Schwierigkeiten mit der Implementation von Bewertungskompetenz ließen sich auf den Ebenen ‚Schule‘, ‚Lehrkraft‘ und ‚Schüler‘ zusammenfassen. Den Auskünften der Lehrkräfte zufolge waren dabei auf der Ebene der Lehrkraft neben einem Mangel an geeigneten Unterrichtsmaterialien zur Förderung von Bewertungskompetenz auch ein Mangel an eigener Kompetenz, entsprechende Unterrichtsstunden oder -einheiten durchführen zu können, als Schwierigkeiten existent. Weitere Problembereiche stellten die Ergebnisoffenheit ethischer Fragestellungen, der Umgang mit der eigenen Meinung sowie die Beurteilung von Schülerleistungen in diesem Bereich dar.

Sowohl in Bezug auf das „deutsche“ Konstrukt der Bewertungskompetenz als auch im Kontext des internationalen Ansatzes der SSI – die ein in Teilen der Bewertungskompe-

¹Das dreidimensionale ESNaS-Modell wurde in dieser Studie nicht weiter berücksichtigt, da es für einen Einsatz in Large-Scale-Assessments angelegt ist und daher für eine Anwendung im Kontext der Diagnose von Schülerleistungen durch Lehrkräfte im Unterricht weniger geeignet scheint.

tenz vergleichbares, im Detail jedoch umfassenderes Konstrukt bilden – bestehen Schwierigkeiten mit der Beurteilung von Schülerleistungen (Fensham und Rennie 2013, S. 80): „[A]s a mainstream intention of the curriculum, ‚decision-making‘ has been dogged by a lack of means of assessing it.“ Sadler (2011) stellt fest, dass die Behandlung von SSI seitens der Lehrkraft Hingabe und die Bereitschaft voraussetzt, mit Unsicherheiten umgehen zu können. Auch Willmott und Willis (2008) weisen auf diese Problematik hin, die sich unter anderem darin begründe, dass Lehrkräfte der Naturwissenschaften stärker die Beurteilung von fachlichem Wissen als die Einordnung argumentativer Strukturen als Bestandteil ihres Unterrichts wahrnehmen. Entsprechend ist das Problem der Diagnose von Schülerleistungen im Zusammenhang mit ethischer Bewertung und Urteilsbildung auch im internationalen Diskurs nicht vollständig gelöst. Diese Ausgangslage aufgreifend untersucht diese Studie explizit den Bereich der Diagnose von Bewertungskompetenz, um die hier zentralen Schwierigkeiten empirisch zu erfassen.

Entwicklung der Fragestellung

Das Erkenntnisinteresse der vorliegenden Arbeit verortete sich zunächst im Feld der diagnostischen Fähigkeiten von Lehrkräften hinsichtlich der Schülerkompetenz Bewertung, da dieser Aspekt in Studien zu Bewertungskompetenz bzw. SSI bislang eher peripher behandelt wurde. Die Relevanz der Fragestellung leitete sich aus den in den vorangegangenen Ausführungen behandelten Themenfeldern und des entsprechenden Forschungsstandes, aber auch aus Gesprächen mit Lehrkräften ab. Anfangs war dabei das Gelingen der Diagnose von Schülerleistungen in Bewertungsprozessen von Interesse. Gleichzeitig erschien jedoch eine evaluative Untersuchungsfrage im Hinblick auf den Erkenntnisgewinn der Studie weniger ertragreich. Im Laufe einer durchgeführten Pilotstudie und der beginnenden Hauptuntersuchung vorliegender Arbeit wurde der Fokus auf das Deuten („Wie verstehen Lehrkräfte Bewertungskompetenz und Diagnose?“) und Handeln der Lehrkräfte („Wie gehen Lehrkräfte mit Bewertungskompetenz und ihrer Diagnose um?“) gelegt. Die aus der Wissenssoziologie stammenden Begriffe der Deutungs- und Handlungsmuster (Sackmann 1992; Arnold 1983) schienen aufgrund ihres umfassenden Charakters geeignete Strukturen zur Erklärung von im Datenmaterial vorhandenen Phänomenen zu liefern. Deutungsmuster sind im Rahmen dieser Studie als beständige, aber dennoch veränderbare Interpretationen des Selbst und der Umwelt von Angehörigen einer bestimmten sozialen Gruppe – im vorliegenden Fall Biologielehrkräfte – definiert, welche das Individuum durch Komplexitätsreduktion und den Rekurs auf verfügbare Interpretationen handlungsfähig machen.

Der Begriff der Deutungsmuster wurde gewählt, da er eine Entsprechung im Bereich des Handelns (Handlungsmuster) besitzt und so die beiden Ebenen einerseits getrennt behandelt, andererseits potentiell wechselseitige Einflüsse ermittelt werden können. Die enger gefassten Begriffe ‚Haltung‘ und ‚Einstellung‘ bieten diese (begriffliche) Option nicht. Handlungsmuster können als Formen des Handelns innerhalb einer Gesellschaft, aus welcher heraus sie entwickelt wurden, verstanden werden (Ehlich und Rehbein 1979). Sie können sich dabei auf Individuen, aber auch auf Gruppen beziehen. Einem Individuum zur Verfügung stehende Muster werden dabei zweckgerichtet und situationsgebunden eingesetzt. Die der vorliegenden Untersuchung zugrunde liegende Forschungsfrage lautet somit:

Über welche Deutungs- und Handlungsmuster verfügen Biologielehrkräfte zur Diagnose von Bewertungskompetenz?

Forschungsdesign und Methodik

Grounded Theory als methodologischer Zugang

In dieser explorativen Studie sollen subjektive Sichtweisen und Deutungen der Befragten zur Diagnose von Bewertungskompetenz rekonstruiert werden. Darüber hinaus soll die Herausarbeitung von Bedingungen und Folgen von Sichtweisen im Fokus stehen. Somit wird für die Beantwortung der Forschungsfrage ein rekonstruktiver Forschungsansatz gewählt. Die Grounded Theory in der Prägung von Strauss und Corbin (1996, 1998) dient als methodologischer Zugang, da sie einen umfassenden, auf die Entwicklung einer in den Daten gegründeten Theorie über einen Untersuchungsgegenstand gerichteten Ansatz darstellt. Daher liegt der Akzent auf einer induktiven – bzw. abduktiven (Reichert 2013) – Vorgehensweise und der Ausarbeitung von Beziehungen zwischen Kategorien.

Die Grounded Theory ist ein kodierendes Verfahren, welche im Ansatz von Strauss und Corbin (1996) die iterativ ablaufenden Prozeduren des offenen, axialen und selektiven Kodierens beinhaltet. Das offene Kodieren dient der Entfaltung von Interpretationsmöglichkeiten durch das Verfassen von Memos und die Zuordnung von verallgemeinernden Konzepten zu bestimmten Textstellen des Datenmaterials (Tab. 1).

Im Verlauf des axialen Kodierens werden bereits vorhandene Konzepte (empirischen Ereignissen zugeordnete, verallgemeinernde Begriffe) zu Konzepten höheren Abstraktionsgrades bzw. zu Kategorien zusammengefasst. Es erfolgt ein In-Beziehung-Setzen der Konzepte mittels des sogenannten paradigmatischen Modells (auch Kodierparadigma genannt), d. h. die Konzepte werden geordnet. Zentrale Phänomene werden auf ihre ursächlichen, kontextuellen und intervenierenden Bedingungen, Handlungsstrategien

Tab. 1 Auszug aus einem Interviewtranskript und dazugehörige offene Kodierungen

Interviewtranskript	Kodierungen
I: Inwiefern fließen Leistungen grundsätzlich in deinem Biounterricht im Bereich Bewertungskompetenz in die Notengebung ein?	
B: Ja, ich habe ja eben schon gesagt, da ich da selbst sehr unsicher bin und ich es teilweise schwierig finde, es zu bewerten, äh, fällt es sehr, fließt es sehr vorsichtig ein. Aber das Beispiel, was ich gerade genannt habe, da ließ sich dann schon zumindest sagen wir mal so ‚Mittelklasse‘, ‚sehr gut‘, ‚echt nix dazugelern‘ oder ‚kann er noch nicht‘, ja? Wir formulieren ja hier immer Potenziale eher [...]. Das ließ sich schon gut machen, das konnte ich dann auch begründet wieder zurückgeben und sagen: ‚Du hast deshalb eine zwei, eine vier oder so.‘ Weil ich muss die Noten ja auch entsprechend objektiv begründen und kann nicht nur ‚gefühlte drei‘ drunterschreiben oder so. Und das ist natürlich schwierig, deshalb bin ich da sehr vorsichtig mit	Unsicherheit Schwierigkeit der Bewertung Feedback
I: Ja. Wie gehst du, wenn du jetzt so eine Dilemmasituation im Unterricht behandelst, mit deiner eigenen Meinung zu diesem Dilemma um?	
B: Mhm, ähm, ich versuche es so wie die Schüler auch, dass ich die schon auch sozusagen soweit darstelle, soweit sie schülerverträglich irgendwie ist. Aber immer auch gleichzeitig in Frage stelle, weil ich sage, es gibt ja auch eine andere Sicht darauf. Oder wenn ich jetzt so und so handeln würde, ähm, dass ich sozusagen mich dann als Möglichkeit darstelle. Aber das Entscheidende ist ja gar nicht, wie ICH mich entscheide, sage ich dann auch, weil es gibt ja nicht das RICHTIGE Entscheiden, nicht dass die wieder denken: ‚Guck, der vorne, wie der sich vorne entscheidet, so ist es richtig!‘, sondern nur vielleicht aus meiner Geschichte heraus. Und ich sage vielleicht, früher hab ich da vielleicht SO drüber nachgedacht, heute sehe ich sowas SO. [...]	Meinungsäußerung Lehrkraft Eigene Meinung in Frage stellen Eigene Meinung als Möglichkeit darstellen Individuelle Gründe für Entscheidungen

und Konsequenzen hin untersucht (vgl. Abbildung 1), so dass mehrere Achsenkategorien entstehen. Diese Bereiche kann man sich wie „Schnitte“ durch das Material vorstellen“ (Strübing 2014, S. 26). Im Rahmen der vorliegenden Untersuchung wurden in diesem Erarbeitungsschritt die drei Achsenkategorien „Bedeutungszuweisung von Bewertungskompetenz“, „Sorge um Autoritätsverlust“ und „Vermeidung“ entwickelt.

Das selektive Kodieren schließlich hat zum Ziel, eine zentrale Kernkategorie zu identifizieren, welche das in Bezug auf die Fragestellung zentrale Phänomen konzeptualisiert. Dazu müssen die zuvor im offenen und insbesondere im axialen Kodieren entwickelten Achsenkategorien, der Struktur des Kodierparadigmas folgend, in eine bereits bestehende oder neu zu entwickelnde Kernkategorie integriert werden. In der vorliegenden Studie wurden die entwickelten Achsenkategorien wieder aufgelöst und in die – im selektiven Kodieren neu erstellte – Kernkategorie „Negiertes Bewältigen“ integriert (Abb. 1). Die Kodierprozeduren wurden in der vorliegenden Studie durch das Programm MAXQDA unterstützt.

Datenerhebung

Um Antworten auf die Fragestellung der Untersuchung zu erlangen, wurden in der vorliegenden Studie Daten in einem insgesamt dreischrittigen Prozess erhoben. In einem problemzentrierten (Witzel 2000) Vor-Interview wurden zunächst Perspektiven und Deutungen der Lehrkräfte im Hinblick auf Bewertungskompetenz und deren Diagnose thematisiert. Das problemzentrierte Interview zielt als ein Verfahren der Datenerhebung auf eine „möglichst unvoreingenommene Erfassung individueller Handlungen

sowie subjektiver Wahrnehmungen und Verarbeitungsweisen gesellschaftlicher Realität“ (Witzel 2000, Abs. 1). Die Konstruktion des problemzentrierten Interviews als Methode lehnt sich explizit an die Prinzipien der Grounded Theory an: Mittels des Leitfadens wird versucht, einerseits dem in der qualitativen Sozialforschung verfolgten Credo der Offenheit, andererseits jedoch gleichzeitig der Forderung nach Theoriegeleitetheit von empirischen Erkenntnissen durch ein deduktiv-induktives Vorgehen zu begegnen (Witzel 2000). So steht insbesondere die in narrativen Ausführungen geschilderte Sicht der Interviewten auf das vom Interviewer fokussierte Problem im Zentrum. Leitfadenbasierte Nachfragen können jedoch in dialogischer Form die Narrationen ergänzen. Wesentliche Auszüge aus dem Leitfaden der Vor-Interviews lauteten:

Was wollen Sie mit Ihrem Biologieunterricht grundsätzlich bei den Schülern erreichen?

Was verstehen Sie persönlich unter dem Begriff „Diagnose“ (im unterrichtlichen Zusammenhang)?

Wie gehen Sie beim Diagnostizieren von Schülerleistungen vor?

In den Bildungsstandards ist der Kompetenzbereich ‚Bewertung‘ festgelegt. Wie würden Sie diesen Kompetenzbereich beschreiben?

Welche Erfahrungen haben Sie bereits mit Bewertungskompetenz im Unterricht gemacht?

Für wie wichtig erachten Sie diesen Kompetenzbereich für Ihr Fach im Vergleich zu den anderen Kompetenzbereichen?

Wie würden Sie einen Schüler beschreiben, der „bewertungskompetent“ ist?

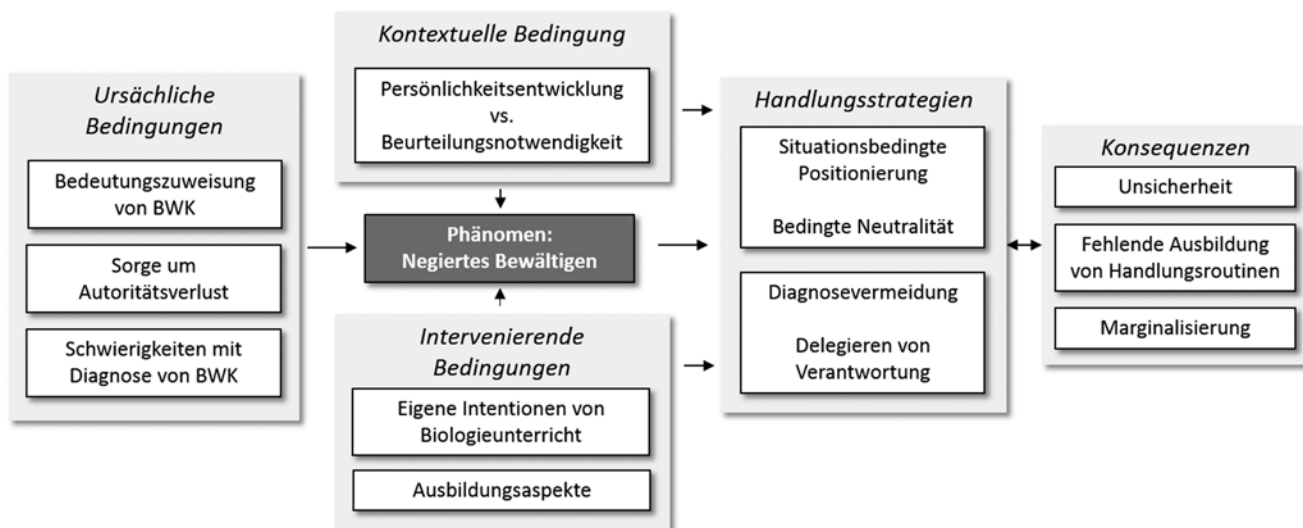


Abb. 1 Modell der Kernkategorie „Negiertes Bewältigen“, dargestellt nach dem Kodierparadigma (Strauss und Corbin 1996; 1998). Das Kodierparadigma verdeutlicht die Beziehungen der einzelnen Konzepte zueinander

Welche Kriterien legen Sie für die Diagnose von Bewertungskompetenz zugrunde?
 Wie gehen Sie mit Ihrer eigenen Meinung zu einem relevanten Problem um?

Fallauswahl

Es folgte die Durchführung und begleitende Videographie eines Rollenspiels im Unterricht der interviewten Lehrkraft. In den Unterrichtssequenzen diskutierten Schüler der Sekundarstufe I eine Dilemmasituation aus dem Themenfeld des Klimawandels. Aus der Videographie wurden in einem dritten Schritt einzelne kurze Videosequenzen (Vignetten; vgl. Brovelli et al. 2013; Rehm und Bölsterli 2014) ausgewählt, welche dann in einem zweiten Interview von der Lehrkraft im Hinblick auf Bewertungskompetenz diagnostisch analysiert wurden. Als Vignetten wurden einzelne Schülerstatements oder kurze Dialoge zweier Schüler gewählt. Auswahlkriterien waren einerseits aus Sicht der Forscher sichtbar gute Leistungen von Schülern im Hinblick auf Bewertungskompetenz, aber auch Ausschnitte, in denen in diesem Zusammenhang Defizite erkennbar waren. Grundlage für die Einordnung der Schüleraussagen waren Teilkompetenzen der Modelle zu Bewertungskompetenz im Biologieunterricht (u. a. Bögeholz 2007; Reitschert et al. 2007).

Im Rahmen von Grounded-Theory-Studien ist das sogenannte theoretische Sampling als Vorgehen bei der Fallauswahl charakteristisch (Strauss und Corbin 1996). Bei diesem Verfahren erfolgen Erhebung und Auswertung nach der Erhebung erster Daten parallel, wobei basierend auf der Analyse des erhobenen Datenmaterials die Entscheidung über weitere einzubeziehende Fälle getroffen wird. Die Erhebungen erfolgen so lange, bis keine neuen relevanten Ähnlichkeiten oder Unterschiede aufgedeckt werden können, die Ergebnisse demnach eine theoretische Sättigung aufweisen (Strauss und Corbin 1996). In dieser Studie wurden nach dem auf der Basis von Vorüberlegungen ausgewählten ersten Fall zur Ausdifferenzierung und Verdichtung der zu entwickelnden Theorieskizze mittels der Methode des ständigen Vergleiches im Hinblick auf verschiedene Dimensionen kontrastierende Fälle ausgewählt (Strauss und Corbin 1996). Kriterien für die Auswahl der Fälle waren sowohl die Dauer der Berufserfahrung (in Relation zur Einführung der Bildungsstandards), das Zweitfach als auch aus dem Datenmaterial generierte Konzepte (s. *Ausgewählte Ergebnisse – Fallkontrastierung im Samplingprozess*). Das

Tab. 2 Übersicht über die im Prozess des theoretischen Samplings in dieser Untersuchung ausgewählten Fälle

Fall	Pseudonym	Fächerkombination	Berufserfahrung (Jahre)	Schulform
1	Herr Degenhardt	Biologie/Sport	9	Gesamtschule
2	Frau Andörfer	Biologie/Chemie	8	Haupt-/Realschule
3	Frau Hofer	Biologie/Chemie/Sport	2	Gesamtschule
4	Frau Vollertsen	Biologie/Politik	8	Gymnasium
5	Herr Bernhold	Biologie/Chemie	21	Gymnasium
6	Frau Rohrbach	Biologie/Politik	20	Gymnasium

finale Sample setzte sich aus sechs Lehrkräften zusammen (Tab. 2).

Ausgewählte Ergebnisse

Im Folgenden sollen ausgewählte Ergebnisse der Studie dargelegt werden. Dazu sollen zunächst die im Prozess des theoretischen Samplings getroffenen Entscheidungen reflektiert werden, um die Fallkontraste zu verdeutlichen. Anhand zweier Beispiele wird anschließend die zentrale Problematik im Kontext der Diagnose von Bewertungskompetenz veranschaulicht, worauf die Präsentation wesentlicher Bestandteile der entwickelten Theorieskizze zur Diagnose von Bewertungskompetenz in Form von Deutungs- und Handlungsmustern erfolgt.

Fallkontrastierung im Samplingprozess

Im Rahmen der vorliegenden Studie wurde zunächst gezielt gesampelt (Strauss und Corbin 1996; Truschkat et al. 2011), indem basierend auf theoretischen Vorüberlegungen zunächst eine Biologielehrkraft interviewt wurde, die eine spezielle Fortbildung in Bezug auf die Förderung von Bewertungskompetenz im Unterricht erfahren hat. Diese Qualifizierung bestand aus einer mindestens einjährigen Teilnahme am Projekt „Der Klimawandel vor Gericht“ (Eilks et al. 2011), in welchem nach dem Prinzip der partizipativen Aktionsforschung (Eilks und Ralle 2002) und in Zusammenarbeit mit verschiedenen Fachdidaktiken die Entwicklung und Erprobung von Unterrichtsmaterialien zur Förderung von Bewertungs- und Urteilskompetenzen im Unterricht erfolgte. Mit dem Aspekt der Fortbildung konnte angenommen werden, dass dieser erste Fall „wichtige Informationen für die Beantwortung der Forschungsfrage“ enthält (Truschkat et al. 2011, S. 362). Ausgehend von diesem Fall einer fortgebildeten Lehrkraft (Herrn Degenhardt) wurde im Sinne einer Verdichtung des vorliegenden Datenmaterials eine weitere fortgebildete Lehrkraft (Frau Andörfer) ausgewählt, die zudem in einer differenten Schulform (Haupt- und Realschule als Kontrast zur Form der Integrierten Gesamtschule der ersten Lehrkraft) tätig war.

Im weiteren Verlauf der Erhebungen wurden zur Ausdifferenzierung und Verdichtung des Datenmaterials im Hinblick auf verschiedene Dimensionen kontrastierende Fälle ausgewählt (Mey und Mruck 2009). Leitend für die Auswahl der dritten Lehrkraft waren die Konzepte „Bewertungskompetenz als Novum“ und „Mangelnde Aus- und Fortbildung“ aus dem ersten erhobenen Fall. Hier machte die entsprechende Lehrkraft deutlich, dass sie Bewertungskompetenz in dieser expliziten Form als etwas Neues für den Biologieunterricht ansähe und sich für die Bewertung

von Schülerleistungen in diesem Bereich bzw. von Kompetenzen im allgemeinen nicht als aus- oder fortgebildet einschätzt. Entsprechend wurde nun also eine Lehrkraft ausgewählt, welche die Einführung der Bildungsstandards für das Fach Biologie bereits als Bestandteil des Studiums wahrgenommen und einen explizit *nicht* an Kompetenzen orientierten Unterricht als aktive Lehrkraft gar nicht erfahren hat. So wurde mit Frau Hofer eine junge Lehrerin ausgewählt, deren Ausbildung entsprechend noch nicht lange zurückliegt. Aus der Auswertung der Interviews mit Frau Hofer resultierte das Konzept „fachübergreifender Charakter von Bewertungskompetenz“. Diese Charakterisierung des untersuchten Kompetenzbereichs war der Grund für die folgende Auswahl einer Politiklehrkraft, die als zweites Fach Biologie unterrichtete (Frau Vollertsen). Da Frau Hofer entsprechend über eine vergleichsweise kurze Zeit der Berufserfahrung verfügte, wurde mit Herrn Bernhold ein Biologielehrer einbezogen, welcher über viel Berufserfahrung und darüber hinaus auch Beteiligungen an der Ausbildung von Lehramtsstudierenden aufwies. Frau Vollertsens Einstieg in den Beruf fiel zeitlich mit der Einführung der bundesdeutschen Bildungsstandards zusammen, so dass sie anmerkte, in ihrem Unterricht zum damaligen Zeitpunkt keinen Umbruch festgestellt zu haben, dieser sich vielmehr erst aktuell zu zeigen beginne. Daraufhin wurde mit Frau Rohrbach eine Politik- und Biologielehrerin ausgewählt, welche bereits deutlich länger im Beruf tätig ist und entsprechend in beiden Fächern auch viele Erfahrungen bereits vor der Zeit der Kompetenzorientierung gemacht hat.

Diagnose von Bewertungskompetenz anhand von Videovignetten: Zwei Fallbeispiele

Beispielhaft sollen an dieser Stelle die Diagnose zweier Lehrkräfte zu Schülerleistungen im Bereich ‚Bewerten‘ angeführt werden. Zunächst sollen als Stimulus in den Nach-Interviews genutzte Videovignetten in transkribierter Form und die nach Aufforderung durch den Interviewer erfolgte Diagnose von Schülerleistungen zu Bewertungskompetenz aus dem betreffenden Nach-Interview beschrieben werden. Anschließend wird das zentrale Phänomen durch weitere Zitate aus den Fällen zugespitzt.

Fallbeispiel I

Vignette

Schülerin: Also, ich finde es eigentlich unsinnig, jeden Tag Fleisch in der Mensa zu verkaufen, weil die Produktion ja auch sehr umweltschädlich ist – wie auch heute noch angesprochen wird – aber auch zuviel Verzehr von Fleisch ist nicht gut für den Körper, aber in Maßen geht es eigentlich

noch. Aber ich bin eigentlich nicht dafür, dass es jeden Tag angeboten wird.

In der Analyse dieser Sequenz nimmt die interviewte Lehrkraft – Herr Degenhardt – zunächst vornehmlich Bezug auf inhaltliche Aspekte der Argumentationen, indem er die vorgebrachten Argumente und deren Begründungen inhaltlich rekapituliert. Dabei kommen jedoch auch Formulierungs- bzw. Artikulationsaspekte in Form der Bestimmtheit von Äußerungen zum Tragen: *„Ich meine, ich habe gehört, sie leitete so ein mit ‚ich finde‘. ‚Ich finde, ich denke, ich glaube‘, müsste man nochmal genau gucken, ob man daran sozusagen was ablesen kann“*. Das Belegen der Aussagen durch Fakten findet ebenfalls in seiner Reflexion Erwähnung: *„Wie weit sie das belegt, wäre ja vielleicht noch eine Qualität. Inwieweit sie deutlich macht, wie weitreichend so ein Argument ist, dass es da um Zukunft geht, was die eigene Gesundheit [...] angeht, um [die] Zukunft der Welt geht, wegen Klimawandel“*. Darauf aufbauend spricht Herr Degenhardt auch die Hinterlegung der Argumente durch ethische Werte an: *„Ob das hinterlegt ist mit diesen Werten sozusagen, ob sie die auch nennt, das habe ich jetzt, glaube ich, erstmal nicht gehört“*. Eine Differenzierung zwischen deskriptiven und normativen Aussagen nimmt Herr Degenhardt hier nicht vor.

Trotz dieser diagnostizierenden Reflexionen verdeutlicht er an anderer Stelle im Interview: *„[Mit der Einführung von Bewertungskompetenz] haben uns die Politiker [...] ja ziemlich alleine gelassen, weil wir eben das nicht diagnostizieren können. Dafür bin ich nicht ausgebildet, dafür bin ich nicht fortgebildet und dafür ist uns auch nichts an die Hand gegeben worden.“*

Fallbeispiel II

Vignette

Schüler: Zum Beispiel Frühkartoffeln aus Ägypten, [...] die verbrauchen so viel Wasser, dass sie schon teilweise schlimmer als biologisch „angebautes“ Fleisch sind. Aber über Kartoffeln regt sich niemand auf, weil Kartoffel ist gesund und gehört auf den Teller. [...] also schlecht „angebautes“ Fleisch aus Massentierhaltung oder so, keine Frage, ist schlecht für die Umwelt. Aber biologisch „angebautes“ Fleisch ist nicht so extrem schlimm, dass man da so eine Panik drum machen muss und außerdem sollte man, wie die Schüler auch schon sagen, dass sie selber entscheiden können, was sie essen. Also es gibt ja vegetarische Möglichkeiten [...] und vielleicht sollte man allgemein das Fleischangebot verringern, dass es nur noch ein oder zwei Tage in der Woche Fleisch gibt, aber das ganz rauszunehmen halte ich für zu radikal.

Die Bewertung dieser Schüleräußerung fällt seitens der interviewten Lehrkraft – Frau Hofer – durchweg positiv

aus: *„Ja, richtig gut (lacht). Also da würde ich sagen, da steckt wirklich alles drin. Also super Argumente, die gegeneinander abgewogen werden, Fachwissen, das irgendwie einbezogen wird, das aus der vorhergehenden Stunde mitgenommen wurde, wichtige Informationen und begründete Argumente und am Ende ja, eine Stellungnahme: ‚Deshalb bin ich der Meinung, dass das und das passieren sollte.‘ So die Folgen jeweils werden auch mit einbezogen, abgewogen. Also das wäre, das wäre für mich glaube ich schon eine sehr gute Leistung.“* Allgemein führt Frau Hofer an anderer Stelle zur Schülerkompetenz Bewertung aus, dass der „bewertungskompetente“ Lernende *„in der Lage ist, wirklich Argumente für seine Aussagen oder seine Urteile vorzutragen, dass er auch die Perspektiven von vielleicht anderen Personen oder von anderen Interessen mit heranziehen kann und die gegeneinander abwägen, um dann zu Entscheidungen zu kommen.“*

Zum Aspekt der Bewertung von Bewertungskompetenz führt sie aus: *„Aber genau, wo ich mir halt unsicher wäre und wo ich auch mir vorstellen könnte halt, dass es [das] für viele Lehrer ist, wenn sie das eben benoten sollen.“* Und weiter: *„Ich glaube [Bewertungskompetenz] zu vermitteln [...], ist deutlich einfacher als zu diagnostizieren oder zu bewerten, ob jetzt ein Schüler Bewertungskompetenz besitzt oder nicht. [...] Also wie soll ich jetzt eigentlich entscheiden oder benoten vielleicht sogar, ob ein Schüler Bewertungskompetenz besitzt oder nicht? Ich glaube, das ist die Sorge, die viele Lehrer haben, dass es einfach schwierig ist, dann das Ganze auch wirklich zu bewerten.“*

Eine Theorieskizze zur Diagnose von Bewertungskompetenz durch Biologielehrkräfte

Die Kernkategorie „Negiertes Bewältigen“

Das Ziel des selektiven Kodierens ist es, die entwickelten Achsenkategorien in ein zentrales Modell – die Kernkategorie – zu integrieren. Diese soll das Schlüsselphänomen im Hinblick auf die Forschungsfrage darstellen (Strauss und Corbin 1996). Die Kernkategorie und die in sie integrierten Achsenkategorien bilden das zentrale Ergebnis – die Grounded Theory – einer gleichnamigen Studie.

Die Erscheinung, die in den Daten dieser Untersuchung immer wieder auftrat und somit im Hinblick auf die Fragestellung als zentral angesehen werden konnte, wurde als „Negiertes Bewältigen“ (Abb. 1) nach dem Kodierparadigma (Strauss und Corbin 1996) ausgearbeitet: Die Lehrkräfte stellten ihre Fähigkeiten zur Diagnose von Bewertungskompetenz entweder in Frage oder stritten sie gar vollständig ab. Darüber hinaus bezeichneten sie sich als unsicher und beklagten einen Mangel an Diagnosekriterien. Jedoch zeigten sie sich trotz allem in der Lage, wesentliche Elemente von Bewertungskompetenz innerhalb der video-

graphierten Schüleraussagen zu diagnostizieren. „Negiertes Bewältigen“ meint somit das Abstreiten, in Frage Stellen und Relativieren eigener nachweislich vorhandener Fähigkeiten. In allen untersuchten Fällen konnten Indikatoren zu diesem Konzept kodiert und die entwickelten Achsenkategorien in die so kreierte Kernkategorie integriert werden. Das zentrale Phänomen konnte durch die im Folgenden dargestellten Konzepte beschrieben werden.

Ursächliche Bedingungen des Phänomens „Negiertes Bewältigen“

Ein in einem sozialen Kontext auftretendes Phänomen ist stets auf bestimmte Ursachen bzw. ursächliche Bedingungen zurückzuführen. Selten ist es dabei der Fall, dass ein solches Phänomen auf nur eine Ursache zurückzuführen ist. Entsprechend ist auch das Kodierparadigma (Strauss und Corbin 1996) als heuristischer Rahmen zur Ausarbeitung von Erscheinungen in den Daten in Form des Fragen-Stellens darauf ausgerichtet, diese ursächlichen Bedingungen eines zentralen Phänomens zu beschreiben

Bedeutungszuweisung von Bewertungskompetenz

Das Konzept „Bedeutungszuweisung von Bewertungskompetenz“ fasst Interviewabschnitte zusammen, die Informationen darüber enthalten, welche Bedeutung die Lehrkräfte dem Kompetenzbereich ‚Bewertung‘ für ihren Biologieunterricht beimessen. Eine hohe Relevanz für das Fach Biologie bescheinigt Herr Degenhardt dem Bereich der Bewertungskompetenz: *„Ich halte sie für sehr wichtig, weil das ein Einfallstor ist, eine Möglichkeit, in diesem Fach wegzukommen von dem enzyklopädischen Wissen.“* Herr Bernhold rückt in seinem Unterricht dagegen – in der Tendenz – inhaltliche Aspekte in den Vordergrund, die jedoch einem bestimmten Zweck dienen sollen: Als zentrales Ziel seines Unterricht beschreibt er *„erstmal natürlich eine Form eines grundlegenden Wissens, das ihnen [den Schülern] die Möglichkeit einräumen soll, später eigene Entscheidungen auf der Basis ihrer Kenntnisse treffen zu können.“* Die Lehrkräfte in dieser Studie wiesen dem Kompetenzbereich ‚Bewertung‘ insgesamt eine hohe Relevanz zu, die jedoch teilweise und spätestens in der unterrichtlichen Praxis offenbar hinter den anderen Kompetenzbereichen zurücktritt.

Sorge um Autoritätsverlust

Als „Sorge um Autoritätsverlust“ wurden Abschnitte kodiert, die Sorgen und Vorbehalte der Lehrkräfte im Hinblick auf Unterrichtssequenzen zum ethischen Bewerten und die damit verbundene Rolle der Lehrkraft thematisieren. Änderungen in der Rolle sind durch Verlagerungen der schwerpunktmäßigen Aktivität der Lehrkräfte bedingt: Ist

sie im „normalen“ Unterricht oftmals eher als lenkend zu beschreiben, kommt im Rahmen der Förderung ethischer Bewertungskompetenz vielmehr eine moderierende Funktion zum Tragen. Die Lehrkraft ist hier nicht Inhaber eines Wissensbestandes, den sich die Lernenden im Unterricht unter ihrer Anleitung aneignen bzw. erarbeiten müssen. *„Mir fällt es immer noch schwer manchmal solche Diskussionen zu organisieren, zu leiten in einem Klassenverband“* beschreibt Herr Degenhardt. Mit dieser Änderung in der Rolle der Lehrkraft geht die Sorge einher, als unterrichtsleitende Instanz an Autorität zu verlieren. Die Sorge um Autoritätsverlust schlägt sich auch im Charakter der anzulegenden Diagnosekriterien nieder, da hier die Gefahr einer für die Schüler nicht nachvollziehbar zu begründenden Leistungsbeurteilung gesehen wird: *„Und du gibst dem einen Schüler eine eins für seine Bewertungskompetenz und dem anderen eine drei und dann musst du ja auch deine Noten begründen können“* (Frau Hofer). Herr Bernhold beschreibt die Lehrkraft im Kontext von Bewertungsprozessen als eine Person, die wie ein *„möglichst unpersönliches – in Anführungsstrichen – alles zulassendes Wesen daherschwebt“*.

Schwierigkeiten mit Diagnose von Bewertungskompetenz

In diesem Konzept wurden Äußerungen gruppiert, die Aufschluss über Probleme und Schwierigkeiten mit der Diagnose von Bewertungskompetenz gaben. Es dominierten hier Aussagen, welche einen Mangel an (geeigneten) Diagnosekriterien thematisierten. Im Hinblick auf Bewertungskriterien zu Bewertungskompetenz bewegen sich die Äußerungen der Lehrkräfte im Rahmen dieser Untersuchung zwischen einem postulierten Mangel und einer in Frage gestellten Geeignetheit potentieller Kriterien im Hinblick auf Sachlichkeit und Objektivität. *„Ja, und dann eben auch, dass so die Kriterien fehlen. Also was ist jetzt wirklich ein sehr guter, bewertungskompetenter Schüler?“*, fragt Frau Hofer.

Im Widerspruch zu diesen Befunden stehen dabei die Vielzahl und die Differenziertheit der von den Biologielehrkräften zur Diagnose von Bewertungskompetenz genannten bzw. zur Diagnose der in den Vignetten sichtbaren Schülerleistungen herangezogenen Kriterien. In den Diagnosen innerhalb der Vignetten wurden u. a. folgende Kriterien verwendet (bzw. an anderer Stelle genannt): Fachliche Richtigkeit von Argumenten, Reichweite von Argumenten, Argumentationsfähigkeit, Eingehen auf Diskussionspartner, Erkennen unterschiedlicher Positionen, Perspektivenwechsel, Nennen und Reflexion von Folgen, Abwägen von Argumenten, Entscheidungsfällung, Nennen von Werten, Nennen von Argumenten, Begründen von Argumenten, Nennen von Argumenten für unterschiedliche Positionen, Spannungstoleranz unterschiedlicher Meinungen, Ernsthaftigkeit der Auseinandersetzung. Viele dieser Kriterien lassen sich den

Teilkompetenzen der Modelle zu Bewertungskompetenz zuordnen.

Kontextuelle Bedingung des Phänomens „Negiertes Bewältigen“

Der Kontext, in dem ein soziales Phänomen auftritt, wird als ein weiterer Bestandteil des paradigmatischen Modells und damit als an das jeweilige Phänomen zu stellende Frage eingeordnet (Strauss und Corbin 1996). „Der Kontext stellt den besonderen Satz von Bedingungen dar, in dem die Handlungs- und interaktionalen Strategien stattfinden“ (Strauss und Corbin 1996, S. 75).

Persönlichkeitsentwicklung vs. Beurteilungsnotwendigkeit

Als „Persönlichkeitsentwicklung vs. Beurteilungsnotwendigkeit“ wurden Sequenzen kodiert, die Informationen über die Persönlichkeitsentwicklung der Lernenden und die Notwendigkeit einer Beurteilung von Schülerleistungen im schulischen Umfeld enthielten. Bewertungskompetenz wird von den Lehrkräften als ein Persönlichkeitsmerkmal angesehen, das zur Zukunftsfähigkeit der Lernenden einen erheblichen Beitrag leisten kann. „Also das ist ja [...] ein Persönlichkeitsmerkmal, auf das ich hinarbeite“, verdeutlicht Herr Bernhold. Die Entwicklung der Persönlichkeit der Lernenden gehört ebenfalls zu den Aufgaben, die die Institution Schule bewerkstelligen soll. Im Kontext der Diagnose von Bewertungskompetenz – also u. a. dem Bewerten von bewertenden Schülern – kommt es zu einem Konflikt zwischen dieser Persönlichkeitsentwicklung und der Beurteilungsnotwendigkeit von Schülerleistungen durch die Lehrkraft: „Und da [...] streiten sich dann bestimmte Aufgabenfelder, die mit Schule zu tun haben. Also so Persönlichkeitsentwicklung und Bildung und gleichzeitig auch Beurteilung, das verträgt sich manchmal nicht“ (Herr Bernhold).

Intervenierende Bedingungen des Phänomens „Negiertes Bewältigen“

Strauss und Corbin (1996, S. 75) verstehen unter intervenierenden Bedingungen „die strukturellen Bedingungen, die auf die Handlungs- und interaktionalen Strategien einwirken, die sich auf ein bestimmtes Phänomen beziehen.“ Entsprechend können Handlungsstrategien durch bestimmte intervenierende Bedingungen erleichtern oder erschwert werden.

Eigene Intentionen von Biologieunterricht

Das Konzept „Eigene Intentionen von Biologieunterricht“ fasst Aussagen der interviewten Lehrkräfte zusammen, die

sich auf allgemeine, übergeordnete Ziele beziehen, welche die jeweilige Lehrkraft mit ihrem Biologieunterricht im Hinblick auf die Lernenden verfolgt. Bei allen interviewten Biologielehrkräften standen die Vermittlung naturwissenschaftlicher Sicht- und Denkweisen und deren Alltagsbezug im Fokus. Stellvertretend äußert dazu Herr Degenhardt: „Also sozusagen auf dieses Fachwissen, auf diese Einzel-fakten nicht zu viel Wert zu legen, sondern auf die Denkweise, auf die Zusammenhänge, auf die Erkenntnis, auf die Strukturen, die man erkennen kann, dass man Gesetzmäßigkeiten erkennt.“ Keine der interviewten Biologielehrkräfte stellte die ausschließliche Konzentration auf Fachwissen in den Fokus ihres Unterrichts. Vielmehr konzentrierten sich ihre Aussagen auf eine Ebene, die das Fachwissen als Ausgangs- und Entscheidungsbasis für alltags- und zukunftsrelevante Fragestellungen heranzieht.

Ausbildungsaspekte

Die von der jeweiligen Lehrkraft genossene Ausbildung sowie besuchte Fortbildungen wirken als eine weitere intervenierende Bedingung auf die angewandten Handlungsstrategien und auf das durch die Kernkategorie dargestellte Phänomen „Negiertes Bewältigen“ ein. Der Zeitpunkt der Ausbildung relativ zur Einführung der Bildungsstandards ist hier als relevant zu erachten. Dieser Aspekt war im Rahmen der vorliegenden Untersuchung auch ein Kriterium bei der Auswahl der Fälle. Herr Degenhardt erlebte die Kompetenzorientierung erst in der zweiten Phase der Ausbildung: „Und dann bin ich ins Referendariat gegangen und da [gab es dann] das erste Mal diese neuen Kompetenzen zum mittleren Abschluss. Und da haben wir dann gesagt: ‚Oh, Bewertungskompetenz gleichberechtigt neben drei anderen Kompetenzen!‘ Das weiß ich noch, das war dann auf einmal der Paradigmenwechsel.“

Handlungsstrategien zum Phänomen „Negiertes Bewältigen“

Handlungsstrategien sind ein wesentlicher Bestandteil von sozialen Phänomenen und sind entsprechend auch in das paradigmatische Modell (Strauss und Corbin 1996; 1998) integriert. Im Falle der vorliegenden Studie geben die Handlungsstrategien Antworten auf die Fragen wie Lehrkräfte mit der Notwendigkeit einer Integration und Diagnose von Bewertungskompetenz umgehen und welche Handlungsstrategien zur Entstehung und Aufrechterhaltung des zentralen Phänomens „Negiertes Bewältigen“ beitragen.

Situationsbedingte Positionierung

Dem Konzept „Situationsbedingte Positionierung“ wurden im Kodierprozess Sequenzen zugeordnet, die den Umgang

mit der Lehrermeinung im Zusammenhang mit der Thematisierung ethischer Dilemmata beschreiben. Im Hinblick auf das Diagnostizieren von Bewertungskompetenz spielt das Äußern der Lehrermeinung insbesondere bezüglich der Gefahr einer möglichen Beeinflussung der Schüler eine Rolle. Den Aussagen der Lehrkräfte erfolgt eine Positionierung nicht als grundsätzliche Ja-oder-Nein-Entscheidung, sondern vielmehr situationsabhängig. *„Meine eigene Meinung halte ich bei PID [Präimplantationsdiagnostik] oder ähnlichem meistens zurück“* (Herr Bernhold). Vorteile einer Äußerung der Lehrermeinung beschreibt Her Degenhardt: *„Also manchmal ist es ja gut, mit einer eigenen Meinung zu provozieren“*.

Bedingte Neutralität

Das Konzept „Bedingte Neutralität“ fasst Aussagen zusammen, die das Verhalten der Lehrkraft gegenüber geäußerten Schülermeinungen und -urteilen zu ethischen Sachverhalten oder Entscheidungssituationen im Kontext von Bewertungsprozessen adressieren. Eine gegebene Begründung vorausgesetzt, gilt es seitens der Lehrkraft, unterschiedliche Urteile zu tolerieren: *„Wenn das nicht meine eigene Position ist, dann – wenn er oder sie die sachgerecht vertritt – muss ich sie ja als solche erstmal stehenlassen können und darf sie nicht mit ‚mangelhaft‘ beurteilen (lacht), das wäre ja ein bisschen eigenartig“*, fasst Herr Bernhold zusammen. Gleichzeitig hat diese Toleranz von Meinungen jedoch auch Grenzen, da die Lehrkraft *„natürlich auch in gewissem Sinne, wenn es jetzt politisch völlig daneben ist, einen gewissen Einfluss nehmen soll“* (Herr Bernhold). Für die Lehrkräfte ist im Zuge einer normativen Offenheit ethischer Bewertungsprozesse wesentlich, dass *„die Meinung im Rahmen des Wertekonsens‘ bleibt“* (Frau Rohrbach). Diese Grenzen müssen durch die Lehrkraft ebenfalls diagnostiziert werden.

Diagnosevermeidung

Als „Diagnosevermeidung“ wurden Interviewabschnitte kodiert, die auf die Vermeidung der Diagnose von Schülerleistungen zum Kompetenzbereich ‚Bewertung‘ verwiesen. So äußert z. B. Herr Degenhardt, dass er Unterrichtseinheiten zu Bewertungskompetenz *„oft so ein bisschen aus dieser Bewertung rausnehme“*, um auf diese Weise die Notwendigkeit einer ‚Bewertung von Bewertungen‘ zu umgehen. Hier wird also das Diagnostizieren entsprechender Schülerleistungen in der Unterrichtspraxis den Aussagen der Lehrkraft nach offenbar nicht selten ganz aus der unterrichtlichen Diagnostizität ausgeklammert, Schülerleistungen also nicht diagnostiziert.

Delegieren von Verantwortung

Als „Delegieren von Verantwortung“ wurden Aussagen gekennzeichnet, die im Zusammenhang mit der Integration und Diagnose von Bewertungskompetenz auf die Verantwortlichkeiten anderer Instanzen verwiesen. Das Delegieren von Verantwortung existiert auf zwei unterschiedlichen Ebenen. Auf der einen Ebene werden ein Mangel an Ausbildung im Hinblick auf Bewertungskompetenz sowie ein Mangel an unterstützenden Strukturen bzw. Materialien beklagt: *„Also ich finde, dass wir in unserer Ausbildung viel zu wenig zur Bewertungskompetenz gemacht haben, wenn es denn so einen Stellenwert in der Schule haben soll“* (Frau Hofer). Auf der zweiten Ebene werden die normativen Vorgaben in Form von Bildungsstandards und Kerncurricula kritisiert, indem ihnen eine zu geringe Aussagekraft und damit eine fehlende Unterstützungsfunktion attestiert werden. Somit können sie der Lehrkraft nicht als Orientierung (auch) zur Diagnose von Schülerleistungen im Kompetenzbereich ‚Bewertung‘ dienen: *„Und da gibt das Kerncurriculum zum Beispiel ja auch nicht wirklich viel her“*, verdeutlicht Frau Hofer.

Konsequenzen des Phänomens „Negiertes Bewältigen“

In soziale Kontexte eingebettete Phänomene und darauf gerichtete Handlungen ziehen stets bestimmte Konsequenzen nach sich. So wird im Rahmen des paradigmatischen Modells (Strauss und Corbin 1996) auch nach den Folgen eines untersuchten Phänomens gefragt.

Unsicherheit

Die Kodierung „Unsicherheit“ fasst Interviewsequenzen zusammen, welche die entsprechende Erscheinung als Folge des Umgangs mit und der Diagnose von Bewertungskompetenz auf Seiten der Lehrkräfte beschreiben. So weist Frau Hofer im Zusammenhang mit dem Diagnostizieren von Bewertungsprozessen explizit auf dieses Phänomen hin: *„Aber genau, wo ich mir halt unsicher wäre und wo ich auch mir vorstellen könnte, dass es [das] für viele Lehrer ist, wenn sie das eben benoten sollen.“*

Fehlende Ausbildung von Handlungsrouninen

Das Konzept „Fehlende Ausbildung von Handlungsrouninen“ beinhaltet Interviewpassagen, welche die (fehlende) Entwicklung von Handlungsrouninen und -schemata im Hinblick auf die Diagnose von Bewertungskompetenz adressieren. Die Behandlung bio- und umweltethischer Thematiken im Biologieunterricht wird von den Lehrkräften unter anderem als erfahrungsabhängig angesehen: *„Ja, also da die Erfahrungen einfach fehlen“* (Frau Hofer), werde der Kom-

petenzbereich ‚Bewertung‘ häufig nicht in den Unterricht integriert. Auch Herr Bernhold verweist auf entsprechende Eindrücke: *„Die Schwierigkeit ist, glaube ich, [...] dass das Zuweisen von Bewertungskompetenz bei Lehrkräften eventuell auch nicht bestimmten Regeln unterliegt, sondern auch einer gewissen Erfahrung zugeordnet ist.“*

Marginalisierung

Passagen, die ein „Ins-Abseits-Stellen“ von Bewertungskompetenz thematisierten, wurden mit dem Konzept „Marginalisierung“ kodiert. Die Förderung von Bewertungskompetenz wird – wie bereits beschrieben – als wichtig für die individuelle Entwicklung der Lernenden angesehen, in der unterrichtlichen Praxis jedoch offensichtlich nicht in entsprechendem Ausmaß berücksichtigt: *„Bewertungskompetenz, wie sie im Sinne des Kerncurriculums gefordert ist, kommt unterrichtlich deutlich zu kurz. Also sie taucht eben am Rande auf, aber [...] die Prozessschritte zur Erlangung oder zum Erwerb einer solchen Kompetenz, die kommen unterrichtlich zu kurz“* (Herr Bernhold). Bewertungskompetenz wird „als Zusatz“ (Herr Degenhardt) für den Biologieunterricht gesehen, was den randständigen Charakter – eine Marginalisierung – des Kompetenzbereichs verdeutlicht.

Diskussion

Einordnung zentraler Ergebnisse

Das zentrale Ergebnis und damit den Erkenntnisgewinn der Untersuchung bildet das Phänomen, dass die befragten Lehrkräfte im Zusammenhang mit der Diagnose von Bewertungskompetenz Unsicherheiten und Schwierigkeiten beschrieben, ihre eigenen Fähigkeiten relativierten oder gar abstritten, jedoch gleichzeitig in der Lage waren, wesentliche Aspekte von Bewertungskompetenz in präsentierten Videovignetten zu erkennen und zu bewerten. Dieses zentrale Handlungsmuster wurde durch den Begriff „Negiertes Bewältigen“ gefasst und mittels des Kodierparadigmas (Strauss und Corbin 1996) als Kernkategorie ausgearbeitet. Die im Titel des vorliegenden Artikels enthaltene Frage kann somit mit einem „Sowohl-als auch“ beantwortet werden, da die Diagnose von Bewertungskompetenz, den vorliegenden Daten folgend, *sowohl* das Negieren eigener Fähigkeiten *als auch* das Bewältigen einer Herausforderung beinhaltet.

Die in Grounded-Theory-Studien zu sozialen Phänomenen entwickelten Konzepte sind in der Regel auf mannigfaltige Weise miteinander verbunden. So zeigte sich, dass der Bereich der Diagnose von Schülerleistungen im Kompetenzbereich ‚Bewertung‘ nicht losgelöst vom breiteren Kontext der allgemeinen Kompetenzorientierung und der

Problematik der Integration von ethischen Thematiken in den Biologieunterricht betrachtet werden kann. Die Ergebnisse der vorliegenden Studie lassen entsprechend nicht nur Aussagen zur Diagnose von Bewertungskompetenz zu, sondern sind deutlich breiter gefasst. Hier kommt die umfassende Ausrichtung der Grounded Theory als Methodologie zur Erfassung sozialer Realitäten zum Tragen: *„We realize that, to understand experience, that experience must be located within and can't be divorced from the larger events in a social [...] framework and therefore these are essential aspects of our analyses“* (Corbin und Strauss 2008, S. 8).

Die identifizierten Deutungs- und Handlungsmuster lassen Schwierigkeiten der Lehrkräfte mit der Integration und Diagnose von Bewertungskompetenz auf unterschiedlichen Ebenen erkennen. Im Vergleich zu Lehrkräften, die die Kompetenzorientierung bereits in der ersten Ausbildungsphase kennengelernt haben, bedeutet die Behandlung ethisch aufgeladener Themen im Unterricht für erfahrene Lehrkräfte potentiell eine größere Umstellung. Damit wird auch die Rolle der ersten und zweiten Phase der Lehrerbildung angesprochen. Trotz normativer Vorgaben ist dieser Bereich auch von den jeweiligen Ausbilderinnen und Ausbildern abhängig und bislang kaum empirisch untersucht (Blömeke et al. 2009). Die Lehrerbildung kann potentiell bereits die eigene Einstellung zu ethisch geprägte Inhalten im Biologieunterricht prägen. Dass die Zuschreibung von Relevanz zu Ethik als Element des naturwissenschaftlichen Unterrichts variiert (Sadler et al. 2006), liegt in der Natur der Sache. Dennoch sollte eine gewisse Bedeutung entsprechender Themen von allen Biologielehrkräften wahrgenommen werden, da eine reine Vermittlung von Fachwissen heute nicht mehr den alleinigen Fokus des Biologieunterrichts bilden kann (Bögeholz et al. 2004). Dennoch finden ethische Reflexionen im Fach Biologie – wenn überhaupt – offenbar eher beiläufig statt (Dittmer 2012), die Förderung von Bewertungskompetenz wird also oft nicht explizit adressiert. Dabei werden auch die normativen Vorgaben als Grund genannt. Somit kommt es zu einem paradoxen Szenario: Bildungsstandards und Kerncurricula schreiben den Kompetenzbereich ‚Bewertung‘ erstmals verpflichtend für den Biologieunterricht an deutschen Schulen vor. Gleichzeitig wird ihnen jedoch attestiert, dass ihre grundsätzliche Ausrichtung eine vollständige Integration in den Biologieunterricht verhindert.

Strategien der Vermeidung in Bezug auf die Thematisierung – nicht nur die Diagnose – von Bewertungs- und Entscheidungsfindungsprozessen im Biologieunterricht konnten wie in der vorliegenden Studie auch von Lewis (2006) festgestellt werden: Im Rahmen eines durch die Studie evaluierten, alternativen Ansatzes naturwissenschaftlichen Unterrichts sollten Bewertungsprozesse („decision making“) einen gesteigerten Anteil am Unterricht innehaben. Unterrichtsstunden, in denen Entscheidungsfindungs-

prozesse im Zentrum standen, waren zumeist unstrukturiert, was auf eine Unsicherheit der Lehrkräfte zurückgeführt wurde. Diese Unsicherheit versuchten diese durch Vermeidungsstrategien zu umgehen, indem Bewertungsprozesse – um die dort fehlende Zeit wissend – an Stundenenden verlegt oder von den Lehrkräften grundsätzlich in die Verantwortlichkeiten anderer Schulfächer verwiesen wurden (Lewis 2006; vgl. auch Ekborg et al. 2013). Auch Alfs (2012) fand Hinweise auf Unsicherheiten bei Lehrkräften im Zusammenhang mit Bewertungskompetenz. In der Untersuchung von Lewis (2006) zeigten die Lehrkräfte ebenfalls Anzeichen von Unsicherheit, wenn die Integration von Urteils- und Bewertungsprozessen in den Biologieunterricht im Fokus stand (s. o.).

Unsicherheiten zeigen sich auch im Umgang mit der eigenen Meinung einer Lehrkraft. Die Äußerungen der in der vorliegenden Studie interviewten Lehrkräfte ließen sich als Handlungsmuster eines situationsbedingten Kundtuns der eigenen Meinung interpretieren. In der Untersuchung von Sadler et al. (2006) äußerten die Lehrkräfte sich vor allem dahingehend, dass sie ihre eigenen Meinungen und Werte im Rahmen von SSI nicht zum Gegenstand der Diskussion im Klassenraum machen, sondern diese eher zurückhalten. Auf ein Insistieren von Schülern hin würden die meisten Lehrkräfte jedoch auch ihre eigene Ansicht darlegen.

Die Förderung von Kompetenzen der Bewertung wird von den Lehrkräften als die Entwicklung eines Persönlichkeitsmerkmals angesehen. Diagnosen bzw. Bewertungen der entsprechenden Kompetenzen sind entsprechend vorsichtig vorzunehmen. Hogrefe et al. (2012, S. 26) folgend, kann eine ‚Bewertung einer Bewertung‘ Auswirkungen auf die Achtung der Schüler nach sich ziehen, da Beurteilungen im schulischen Kontext stets die Frage nach ihrer Rezeption auf Seiten der Adressaten aufwerfen: Werden sie von den Schülern als „eine rein ‚in der Sache‘ begründete Beurteilung oder eben auch als Infragestellung mit personalen und möglicherweise sogar moralischen Anteilen“ verstanden?

Die identifizierten Schwierigkeiten im Zusammenhangsmodell „Negiertes Bewältigen“ deuten darauf hin, dass der Kompetenzbereich ‚Bewertung‘ als ein äußerst komplexer Bereich gelten muss, der auch und gerade für Lehrkräfte viele Herausforderungen birgt. Dieser Eindruck wird durch die wenigen bislang existenten nationalen und internationalen Studien zu diesem Thema bestärkt (Lewis 2006; Sadler 2011; Alfs 2012; Fensham und Rennie 2013). Gleichzeitig ist festzustellen, dass auch die Diagnose von Schülerleistungen im Allgemeinen als sehr vielschichtiger Aspekt anzusehen ist, die Diagnose von Bewertungskompetenz somit als „doppelte Schwierigkeit“ gewertet werden kann. Wird die Thematisierung ethischer Dilemmasituationen als eine Fachspezifität aufgefasst, wie empirische Studien nahelegen (Hostenbach und Walpuski 2013), so ist für die Durchführung des Unterrichts das entsprechende fachdidaktische

Wissen notwendig (vgl. Baumert und Kunter 2006; Blömeke et al. 2009). Die in dieser Untersuchung interviewten Lehrkräfte zeigten sich in der Lage, den Kompetenzbereich ‚Bewertung‘ zu beschreiben und entsprechende Schülerleistungen in Videovignetten zu diagnostizieren. Gleichzeitig schränkten sie ihre eigenen Fähigkeiten jedoch ein oder stritten sie vollständig ab. Dieser Widerspruch kennzeichnet das zentrale Phänomen in den Daten. Dem vergleichsweise ausgeprägten Faktenwissen zu Bewertungskompetenz stehen jedoch in den Auskünften der Lehrkräfte zu identifizierende defensive Handlungsmuster entgegen (z. B. Diagnosevermeidung, Delegieren von Verantwortung). Eine vergleichbare negative Rückkopplung stellte Alfs (2012) im Hinblick auf die Integration von Bewertungskompetenz in den Biologieunterricht fest: Sie merkte an, dass die in ihrer Untersuchung interviewten Lehrkräfte über ein fundiertes Wissen über Bewertungskompetenz verfügen, dieses Wissen jedoch nicht in der unterrichtlichen Praxis umgesetzt werde. Dadurch wurde auch hier die Entwicklung von Handlungs-routinen stark eingeschränkt. Diese Befunde werden durch die Ergebnisse der vorliegenden Studie gestützt.

Die Zieldomäne der vorliegenden Studie wird durch bestimmte Dimensionen des fachdidaktischen Wissens als Teil des Professionswissens von Lehrkräften (Park und Oliver 2008) gebildet. Dieses Wissen kann – z. B. in der Lehrerbildung – grundsätzlich vermittelt, aber auch durch gezielte Anregungen und Impulse weiterentwickelt und verändert werden (van Driel et al. 1998). Die vorhandene Basis in Form des fachdidaktischen Wissens von Biologielehrkräften zur Diagnose von Bewertungskompetenz kann aufgrund der aus den Ergebnissen der vorliegenden Untersuchung entwickelten Leitlinien weiter ausgebaut werden und potentiell die Basis für die Ausbildung von Handlungsstrategien in der Unterrichtspraxis bilden. So kann potentiell langfristig eine explizite Diagnose von Bewertungskompetenz erfolgen, welche einer gezielten Förderung von Lernenden im Kompetenzerwerb vorausgehen sollte (Hesse 2014). Unterrichtsbegleitende Diagnosen müssen nicht immer vollständig akkurat sein, solange sich die Lehrkraft der möglichen und wahrscheinlichen Unzulänglichkeit der beurteilenden Aussagen bewusst ist (Hesse 2014). Dennoch können möglichst adäquate Diagnosen eine gezielte Förderung der Lernenden im Erwerb von Bewertungskompetenz im Biologieunterricht effektiv unterstützen und somit die Fähigkeit der Lernenden, Themen der modernen Biologie wie gentechnisch veränderte Lebensmittel, Präimplantationsdiagnostik, alternative Formen der Energiegewinnung oder Eingriffe des Menschen in Ökosysteme bewerten zu können, gezielt stärken.

Grenzen der Untersuchung

Wie in jeder (qualitativen) Studie müssen die Grenzen der vorliegenden Untersuchung bzw. der Aussagekraft ihrer Ergebnisse bestimmt und reflektiert werden. Steinke (2010) bezeichnet die „Limitation“ als eigenständiges Gütekriterium für Studien der qualitativen Sozialforschung, da die Ergebnisse entsprechender Untersuchungen aufgrund in aller Regel nicht gegebener – aber auch nicht angestrebter – repräsentativer Stichproben nicht ohne Weiteres für bestimmte Populationen generalisiert werden können. Folgende Punkte müssen im Hinblick auf die vorliegende Studie unter dem Aspekt der Limitation berücksichtigt werden:

Theoretische Sättigung. Die im Rahmen der Studie untersuchte Anzahl von sechs Fällen muss als auch für qualitative Ansätze vergleichsweise niedrig gelten. Das aufwändige Design schränkte die Einbeziehung weiterer Fälle jedoch stark ein. Entsprechend sollte – dem Vorschlag von Breuer (1999) folgend – in Bezug auf das Ergebnis der Studie (der Grounded Theory) vielmehr von einer „Theorieskizze“, denn von einer „Theorie“ gesprochen werden, da eine Sättigung vermutlich nicht für alle Bereiche der Ergebnisse gleichermaßen beansprucht werden kann. Die vorliegende Untersuchung kann – und will – keine Aussagen über die Häufigkeitsverteilung der Deutungs- und Handlungsmuster innerhalb der Gesamtpopulation z. B. bundesdeutscher Biologielehrkräfte treffen. Quantitative Studien könnten hier entsprechende Hinweise liefern.

Zusammensetzung des Samples. Hinsichtlich der Zusammensetzung des Samples ist zu vermuten, dass durch den Fokus des Projekts auf den Kompetenzbereich ‚Bewertung‘ zumindest im Fach Biologie tendenziell Lehrkräfte gewonnen werden konnten, die diesen Kompetenzbereich als ein wichtiges Element ihres Unterrichts betrachten.

Interviewsituation vs. Unterrichtssituation. Die Diagnosen der Lehrkräfte bezüglich der in den Vignetten sichtbaren Schüleraktivitäten wurden in einer „künstlichen“ (Interview-) Situation erbracht (vgl. Lindmeier 2011; Ruhrig und Höttecke 2015). Die Studie kann somit Aussagen über die grundsätzliche Fähigkeit der Lehrkräfte treffen, in einer solchen herausgehobenen Situation wesentliche Aspekte der Schülerkompetenz ‚Bewertung‘ diagnostizieren zu können. Dies kann wiederum als Indikator für das Vorgehen und die Fähigkeit der Lehrkräfte dienen, im unterrichtlichen Alltag entsprechende Leistungen bewerten zu können. Das tatsächliche Vorgehen im Unterricht wurde in dieser Studie jedoch nicht untersucht. Zu berücksichtigen ist in diesem Zusammenhang eine potentielle Situationspezifität diagnostischer Handlungen, so dass sich zwischen den Reflexionen in der Interviewsituation und dem tatsächlichen Handeln im Unterricht Diskrepanzen ergeben können.

Unterschiedlichkeit der Vignetten. Die Tatsache, dass allen Lehrkräften unterschiedliche Vignetten präsentiert

wurden, schränkt die direkte Vergleichbarkeit der Diagnosen ein. Die Vorgehensweise bot jedoch den Vorzug, dass in der Untersuchung situationsspezifisch an konkreten Deutungen und Handlungen angesetzt werden konnte.

Kenntnis der Modelle zur Bewertungskompetenz. Als erste zwei Fälle wurden Lehrkräfte ausgewählt, von denen bekannt war, dass sie im Hinblick auf Bewertungskompetenz fortgebildet sind. Die Auswahl dieser Fälle erfolgte in der Annahme, dass die Daten aus den Interviews produktiv im Hinblick auf die Fragestellung der Untersuchung sind. Gleichwohl ist aus theoretischer Perspektive zu berücksichtigen, dass diese Lehrkräfte andere Voraussetzungen in die Interviewsituation mitbrachten (z. B. die Kenntnis eines der Modelle).

Didaktische Implikationen

Ebenso wie die naturwissenschaftsdidaktische Forschung zur Optimierung von Lehr-Lernprozessen beitragen will (Schecker et al. 2014), ist es auch das Ziel einer Grounded-Theory-Studie, einen praktischen Nutzen der entwickelten Theorie für die im Untersuchungsfeld aktiven Subjekte aufzuzeigen (Strübing 2014).

Die nach dem paradigmatischen Modell (Strauss und Corbin 1996) ausgearbeitete Kernkategorie „Negiertes Bewältigen“ wurde der Entwicklung von Leitlinien für die Lehrerbildung zugrunde gelegt; sie kann jedoch auch aktiven Lehrkräften zur Bewusstmachung von Schwierigkeiten und Problemen dienen und somit die Entwicklung geeigneter Handlungsstrategien zur Diagnose unterstützen. Theorieskizze und Leitlinien können somit dazu beitragen, dass diagnostische Fähigkeiten auf Seiten der Lehrkräfte weiterentwickelt werden, damit der Aspekt des „Bewältigens“ gestärkt, der des „Negierens“ dagegen reduziert werden kann.

Folgende didaktische Leitlinien für die Lehrerbildung wurden aus den Ergebnissen der Studie abgeleitet:

Eigene Intentionen von Biologieunterricht thematisieren. Welchen Stellenwert Lehrkräfte dem Kompetenzbereich ‚Bewertung‘ in ihrem Biologieunterricht zuschreiben, hängt auch mit ihrem grundsätzlichen Verständnis von Biologieunterricht zusammen. Im Kern dieses Aspekts stehen die zentralen Intentionen, welche die jeweilige Lehrkraft mit ihrem Unterricht verfolgt. Das Verständnis und die Intentionen sind dabei Konstrukte, die sich im Laufe der beruflichen Lehrtätigkeit entwickeln müssen – und vermutlich auch verändern können. Für die Ausbildung von Lehrkräften in der ersten und zweiten Phase kann es also hilfreich sein, ein solches Verständnis und zentrale Intentionen zu adressieren, um ein Bewusstsein für diese zu entwickeln. Im Kontext dieses Verständnisses ist es im Rekurs auf die normativen Vorgaben notwendig, auch dem Kompetenzbereich ‚Bewer-

tung' eine Bedeutung für den eigenen Unterricht und das Lernen der Schüler zuzuweisen.

Spannungsfeld der Diagnose von Bewertungskompetenz bewusst machen. Im Rahmen der Diagnose von Bewertungskompetenz kann ein Bewusstsein dafür hilfreich sein, dass das Diagnostizieren von Bewertung einerseits in eine generelle Beurteilungsnotwendigkeit von Schülerleistungen fällt, die einer gezielten Förderung im Kompetenzerwerb vorausgehen sollte. Andererseits sollte wahrgenommen werden, dass Bewertungskompetenz als Merkmal eines mündigen Bürgers zur Entwicklung der individuellen Persönlichkeit der Schüler beiträgt und sich entsprechend ‚Bewertungen von Bewertungen‘ auch potentiell auf die vom Schüler rezipierte Wahrnehmung und Akzeptanz der eigenen Person durch die Lehrkraft und damit auf die Persönlichkeitsentwicklung der Schüler auswirken können. Ein Bewusstsein für diese beiden, sich im Kontext der Diagnose von Bewertungskompetenz gegenüber stehenden Aufgabenfelder, kann potentiell zur Entwicklung von Handlungsschemata zum ‚Bewerten des Bewertens‘ verhelfen.

Den „Autoritätsverlust“ annehmen. Die Veränderungen in der Rolle der Lehrkraft im Kontext von Diskussionen um ethische Dilemmata manifestieren sich primär im Wegfall eines einheitlichen, inhaltlichen „Unterrichtsziels“, auf welches die Lehrkraft steuert und im damit verbundenen Verlust eines Wissens, das die Lehrkraft besitzt und welches die Lernenden erarbeiten müssen. Vielmehr geht es in diesen Situationen um das Moderieren von Diskussionen, in denen unterschiedliche Ansichten und Perspektiven koexistieren und kein gemeinsamer Konsens gefunden werden muss und soll. Auf Seiten der Lehrkraft erfordert dies, dass die Veränderungen in der Lehrerrolle in diesen Zusammenhängen den Schülern gegenüber bewusst gemacht und von der Lehrkraft selbst angenommen und akzeptiert werden müssen.

Fähigkeiten zur Diagnose von Bewertungskompetenz bewusst machen. Im Rahmen dieser Untersuchung zeigte sich durch die Analyse von Videovignetten in den Interviews, dass die Lehrkräfte trotz aller beschriebenen Schwierigkeiten und Unsicherheiten grundsätzlich in der Lage waren, Aspekte von Bewertungskompetenz in Schüleräußerungen zu erkennen und zu bewerten. Auch wenn diese Diagnosen nicht immer alle Details aufgriffen, konnten sie doch überwiegend als weitestgehend adäquate Bewertungen von mündlichen Schülerbeiträgen ausgelegt werden. Basierend auf ihren diagnostischen Fähigkeiten als Lehrkräfte und der Kenntnis von Bewertungskompetenz konnte diese also in den Vignetten durchaus identifiziert werden. An dieser Stelle könnte es entsprechend helfen, Lehrkräften ihre grundsätzlichen Fähigkeiten zur Diagnose von Bewertungskompetenz bewusst zu machen, um ein Zutrauen in die eigenen Fähigkeiten zu entwickeln. Auf diese Weise können wiederum Handlungsschemata entwickelt werden.

Danksagung Die Studie entstand im Rahmen des Stipendienprogramms der Deutschen Bundesstiftung Umwelt (DBU). Für die Finanzierung des Vorhabens sei der DBU herzlich gedankt.

Open Access Dieser Artikel wird unter der Creative Commons Namensnennung 4.0 International Lizenz (<http://creativecommons.org/licenses/by/4.0/deed.de>) veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden.

Literatur

- Abs, H. J. (2006). Zur Bildung diagnostischer Kompetenz in der zweiten Phase der Lehrerbildung. In C. Allemann-Ghionda & E. Terhart (Hrsg.), *Kompetenz und Kompetenzentwicklung von Lehrerinnen und Lehrern: Ausbildung und Beruf* (S. 217–234). Weinheim: Beltz.
- Alfs, N. (2012). *Ethisches Bewerten fördern. Eine qualitative Untersuchung zum fachdidaktischen Wissen von Biologielehrkräften zum Kompetenzbereich „Bewertung“*. Hamburg: Dr. Kovač.
- Arnold, R. (1983). Deutungsmuster. Zu den Bedeutungselementen sowie den theoretischen und methodologischen Bezügen eines Begriffs. *Zeitschrift für Pädagogik*, 29, 893–912.
- Artelt, C., & Gräsel, C. (2009). Diagnostische Kompetenz von Lehrkräften – Gasteditorial. *Zeitschrift für pädagogische Psychologie*, 23(3–4), 157–160.
- Baumert, J., & Kunter, M. (2006). Stichwort: professionelle Kompetenz von Lehrkräften. *Zeitschrift für Erziehungswissenschaften*, 9(4), 469–520.
- Blömeke, S., Felbrich, A., & Müller, C. (2008). Theoretischer Rahmen und Untersuchungsdesign. In S. Blömeke, G. Kaiser & R. Lehmann (Hrsg.), *Professionelle Kompetenz angehender Lehrerinnen und Lehrer: Wissen, Überzeugungen und Lerngelegenheiten deutscher Mathematikstudierender und -referendare* (S. 15–48). Münster: Waxmann.
- Blömeke, S., Kaiser, G., Lehmann, R., König, J., Döhrmann, M., Buchholtz, C., et al. (2009). TEDS-M: Messung von Lehrerkompetenzen im internationalen Vergleich. In O. Zlatkin-Troitschanskaia, K. Beck, D. Sembill, R. Nikkolas & R. Mulder (Hrsg.), *Lehrprofessionalität. Bedingungen, Genese, Wirkungen und ihre Messung* (S. 181–210). Weinheim: Beltz.
- Bögeholz, S. (2007). Bewertungskompetenz für systematisches Entscheiden in komplexen Gestaltungssituationen Nachhaltiger Entwicklung. In D. Krüger & J. Vogt (Hrsg.), *Theorien in der biologie-didaktischen Forschung. Ein Handbuch für Lehramtsstudenten und Doktoranden* (S. 209–220). Berlin: Springer.
- Bögeholz, S., Höhle, C., Langlet, J., Sander, E., & Schlüter, K. (2004). Bewerten – Urteilen – Entscheiden im biologischen Kontext: Modelle in der Biologiedidaktik. *Zeitschrift für Didaktik der Naturwissenschaften*, 10, 89–115.
- Breuer, F. (1999). *Abseits. Marginale Personen – Prekäre Identitäten*. Münster: LIT.
- Brovelli, D., Bölsterli, K., Rehm, M., & Wilhelm, M. (2013). Erfassen professioneller Kompetenzen für den naturwissenschaftlichen Unterricht: Ein Vignettest mit authentisch komplexen Unterrichtssituationen und offenem Antwortformat. *Unterrichtswissenschaft*, 41, 306–329.
- Corbin, J., & Strauss, A. (2008). *Basics of qualitative research. Techniques and procedures for developing grounded theory*. Thousand Oaks: Sage.

- Dittmer, A. (2012). Wenn die Frage nach dem Wesen des Fachs nicht zum Wesen des Fachs gehört. *Zeitschrift für interpretative Schul- und Unterrichtsforschung*, 1, 146–160.
- Dittmer, A., & Gebhard, U. (2012). Stichwort Bewertungskompetenz. Ethik im naturwissenschaftlichen Unterricht aus sozial-intuitionistischer Perspektive. *Zeitschrift für Didaktik der Naturwissenschaften*, 18, 81–98.
- van Driel, J. H., Verloop, N., & de Vos, W. (1998). Developing science teachers' pedagogical content knowledge. *Journal of Research in Science Teaching*, 35, 673–695.
- Ehlich, K., & Rehbein, J. (1979). Sprachliche Handlungsmuster. In H.-G. Soeffner (Hrsg.), *Interpretative Verfahren in den Sozial- und Textwissenschaften* (S. 243–274). Stuttgart: Metzler.
- Eilks, I., & Ralle, B. (2002). Partizipative fachdidaktische Aktionsforschung – ein Modell für eine praxisnahe curriculare Entwicklungsforschung in der Chemiedidaktik. *Chemie konkret*, 9, 13–18.
- Eilks, I., Feierabend, T., Höble, C., Höttecke, D., Menthe, J., Mrochen, M., & Oelgeklaus, H. (Hrsg.), (2011). *Der Klimawandel vor Gericht. Materialien für den Fach- und Projektunterricht*. Freising: Aulis.
- Ekborg, M., Ottander, C., & Silfver, E. (2013). Teachers' experience of working with socio-scientific issues: a large scale and in depth study. *Research in Science Education*, 43, 599–617.
- Fensham, P. J., & Rennie, L. J. (2013). Towards an authentically assessed science curriculum. In D. Corrigan, R. Gunstone & A. Jones (Hrsg.), *Valuing assessment in science education: Pedagogy, curriculum, policy* (S. 69–100). Dordrecht: Springer.
- Hartmann-Mrochen, M. (2013). *Zwischen Notengebung und Urteilsfähigkeit. Einstellungen und Vorstellungen von Lehrkräften verschiedener Fachkulturen zum Kompetenzbereich Bewertung der Nationalen Bildungsstandards*. Dissertation Universität Hamburg, 2012. <http://ediss.sub.uni-hamburg.de/volltexte/2013/6223/pdf/Dissertation.pdf>. Zugriffen: 05. Juli 2015.
- Helmke, A. (2009). *Unterrichtsqualität und Lehrerprofessionalität. Diagnose, Evaluation und Verbesserung des Unterrichts*. Seelze-Velber: Kallmeyer-Klett.
- Hesse, I. (2014). Pädagogisch-psychologische Diagnostik für Lehrkräfte – Herausforderung, Aufgaben, Probleme. In A. Fischer, C. Höble, S. Jahnke-Klein, H. Kiper, M. Komorek, J. Michaelis & J. Sjuts (Hrsg.), *Diagnostik für lernwirksamen Unterricht* (S. 15–39). Baltmannsweiler: Schneider Hohengehren.
- Hesse, I., & Latzko, B. (2011). *Diagnostik für Lehrkräfte*. Opladen: Barbara Budrich.
- Hogrefe, J., Hollstein, O., Meseth, W., & Proske, M. (2012). Die Kommunikation von Urteilen im Unterricht. Zwischen der Bildung und der Beurteilung von Urteilen und deren Folgen. *Zeitschrift für interpretative Schul- und Unterrichtsforschung*, 1, 7–30.
- Hostenbach, J., & Walpuski, M. (2013). Untersuchung der Einflussfaktoren auf die Bewertungskompetenz im Fach Chemie. *Zeitschrift für Didaktik der Naturwissenschaften*, 19, 129–157.
- Hostenbach, J., Fischer, H. E., Kauertz, A., Mayer, J., Sumfleth, E., & Walpuski, M. (2011). Modellierung der Bewertungskompetenz in den Naturwissenschaften zur Evaluation der Nationalen Bildungsstandards. *Zeitschrift für Didaktik der Naturwissenschaften*, 17, 261–288.
- Jones, A., McKim, A., & Reiss, M. (2010). Towards introducing ethical thinking in the classroom: Beyond rhetoric. In A. Jones, A. McKim & M. Reiss (Hrsg.), *Ethics in the science and technology classroom: A new approach to teaching and learning* (S. 1–5). Rotterdam: Sense.
- Karst, K. (2012). *Kompetenzmodellierung des diagnostischen Urteils von Grundschullehrkräften*. Münster: Waxmann.
- Kultusministerkonferenz (KMK). (2005). *Bildungsstandards im Fach Biologie für den mittleren Schulabschluss. Beschluss vom 16.12.2004*. München: Luchterhand.
- Lewis, J. (2006). Bringing the real world into the biology curriculum. *Journal of Biological Education*, 40, 101–106.
- Lindmeier, A. (2011). *Modeling and measuring knowledge and competencies of teachers. A threefold domain-specific structure model for mathematics*. Münster: Waxmann.
- Mey, G., & Mruck, K. (2009). Methodologie und Methodik der Grounded Theory. In W. Kempf & M. Kiefer (Hrsg.), *Forschungsmethoden der Psychologie. Zwischen naturwissenschaftlichem Experiment und sozialwissenschaftlicher Hermeneutik* (Band 3, S. 100–152). Berlin: Regener.
- Mrochen, M., & Höttecke, D. (2012). Einstellungen und Vorstellungen von Lehrpersonen zum Kompetenzbereich Bewertung der Nationalen Bildungsstandards. *Zeitschrift für interpretative Schul- und Unterrichtsforschung*, 1, 113–145.
- Park, S., & Oliver, J. S. (2008). Revisiting the conceptualisation of pedagogical content knowledge (PCK): PCK as a conceptual tool to understand teachers as professionals. *Research in Science Education*, 38, 264–284.
- Ratcliffe, M., & Grace, M. (2003). *Science education for citizenship. Teaching socio-scientific issues*. Maidenhead: Open University Press.
- Rehm, M., & Bölsterli, K. (2014). Entwicklung von Unterrichtsvignetten. In D. Krüger, I. Parchmann & H. Schecker (Hrsg.), *Methoden in der naturwissenschaftsdidaktischen Forschung* (S. 213–225). Berlin: Springer.
- Reichert, J. (2013). *Die Abduktion in der qualitativen Sozialforschung. Über die Entdeckung des Neuen*. Wiesbaden: VS.
- Reitschert, K., Langlet, J., Höble, C., Mittelsten Scheid, N., & Schlüter, K. (2007). Dimensionen ethischer Urteilskompetenz – Dimensionierung und Niveaunkretisierung. *Der mathematische und naturwissenschaftliche Unterricht*, 60(1), 43–51.
- Ruhrig, J., & Höttecke, D. (2015). Components of science teachers' professional competence and their orientational frameworks when dealing with uncertain evidence in science teaching. *International Journal of Science and Mathematics Education*, 13, 447–465.
- Sackmann, R. (1992). Das Deutungsmuster „Generation“. In M. Meuser & R. Sackmann (Hrsg.), *Analyse sozialer Deutungsmuster. Beiträge zur empirischen Wissenssoziologie* (S. 199–216). Pfaffenweiler: Centaurus.
- Sadler, T. D. (2011). Socio-scientific issues-based education: What we know about science education in the context of SSI. In T. D. Sadler (Hrsg.), *Socio-scientific issues in the classroom. Teaching, learning and research* (S. 355–369). Dordrecht: Springer.
- Sadler, T. D., Amirshokohi, A., Kazempour, M., & Allspaw, K. M. (2006). Socioscience and ethics in science classrooms: Teacher perspectives and strategies. *Journal of Research in Science Teaching*, 43, 353–376.
- Schecker, H., Parchmann, I., & Krüger, D. (2014). Formate und Methoden naturwissenschaftsdidaktischer Forschung. In D. Krüger, I. Parchmann & H. Schecker (Hrsg.), *Methoden in der naturwissenschaftsdidaktischen Forschung* (S. 1–15). Berlin: Springer.
- Shulman, L. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, 57, 1–22.
- Steffen, B., & Höble, C. (2014). Decision-making competence in biology education: Implementation into German curricula in relation to international approaches. *Eurasia Journal of Mathematics, Science & Technology Education*, 10, 343–355.
- Steinke, I. (2010). Gütekriterien qualitativer Forschung. In U. Flick, E. von Kardorff & I. Steinke (Hrsg.), *Qualitative Forschung. Ein Handbuch* (S. 319–331). Reinbek bei Hamburg: Rowohlt.
- Strauss, A., & Corbin, J. (1996). *Grounded Theory: Grundlagen qualitativer Sozialforschung*. Weinheim: Beltz. Psychologie Verlagsgesellschaft.

- Strauss, A., & Corbin, J. (1998). *Basics of qualitative research. Techniques and procedures for developing grounded theory*. Thousand Oaks: Sage.
- Strübing, J. (2014). *Grounded Theory. Zur sozialtheoretischen und epistemologischen Fundierung eines pragmatistischen Forschungsstils*. Wiesbaden: VS.
- Tepner, O., Borowski, A., Dollny, S., Fischer, H. E., Jüttner, M., Kirschner, S., & Wirth, J. (2012). Modell zur Entwicklung von Testitems zur Erfassung des Professionswissens von Lehrkräften in den *Naturwissenschaften. Zeitschrift für Didaktik der Naturwissenschaften*, 18, 7–28.
- Truschkat, I., Kaiser-Belz, M., & Volkmann, V. (2011). Theoretisches Sampling in Qualifikationsarbeiten: Die Grounded-Theory-Methodologie zwischen Programmatik und Forschungspraxis. In G. Mey & K. Mruck (Hrsg.), *Grounded Theory Reader* (S. 353–379). Wiesbaden: VS.
- Weinert, F. E. (2001). Vergleichende Leistungsmessung in Schulen – eine umstrittene Selbstverständlichkeit. In F. E. Weinert (Hrsg.), *Leistungsmessung in Schulen* (S. 17–31). Weinheim: Beltz.
- Willmott, C., & Willis, D. (2008). The increasing significance of ethics in the bioscience curriculum. *Journal of Biological Education*, 13, 181–208.
- Witzel, A. (2000). Das problemzentrierte Interview. *Forum Qualitative Sozialforschung* 1, No.1, Art. 22. <http://www.qualitative-research.net/index.php/fqs/article/view/1132/2519>. Zugegriffen: 06.11.2013.