

SABINA EGGERT, SUSANNE BÖGEHOLZ, RAINER WATERMANN UND MARCUS HASSELHORN

Förderung von Bewertungskompetenz im Biologieunterricht durch zusätzliche metakognitive Strukturierungshilfen beim Kooperativen Lernen – Ein Beispiel für Veränderungsmessung

The Effects of Metacognitive Instruction on Students' Socioscientific Decision Making – An Exemplary Procedure for Measurement of Change

Zusammenfassung

In diesem Beitrag wird eine Interventionsstudie zur Förderung von Bewertungskompetenz bei Schülerinnen und Schülern im Biologieunterricht vorgestellt. In der Studie wurden zwei Treatmentvarianten Kooperativen Lernens entwickelt und jeweils in einer Unterrichtseinheit zum Thema Fließgewässerbewertung realisiert. Die Treatmentvarianten unterscheiden sich in der Ab- bzw. Anwesenheit einer zusätzlichen metakognitiven Strukturierungshilfe. An der Untersuchung nahmen 258 Schülerinnen und Schüler der siebten Jahrgangsstufe teil. In einem quasi-experimentellen Prä-Posttest-Design wurde der Zuwachs an Bewertungskompetenz gemessen. Vorab wurde sichergestellt, dass Bewertungskompetenz zu beiden Messzeitpunkten ohne Validitätseinschränkungen mit dem gleichen Messinstrument gemessen und mit dem Rasch Partial Credit Modell analysiert werden kann. Die Auswertungen sind exemplarisch dafür, wie Veränderungsmessungen auf Basis von IRT-Modellen erfolgen können. Es konnte ein interventionsbedingter Zuwachs an genereller Bewertungskompetenz nachgewiesen werden. Eine größere Wirksamkeit des Treatments mit metakognitiver Strukturierungshilfe konnte jedoch nicht gezeigt werden. Abschließend werden die Stärken und Grenzen der vorgestellten Interventionsstudie diskutiert.

Schlüsselwörter: Bewertungskompetenz, Kooperatives Lernen, metakognitive Strukturierungshilfen, Veränderungsmessung

Abstract

The presented intervention study describes the effects of two different cooperative learning environments on the acquisition of students' decision making competence about socioscientific issues. The two treatments differed with respect to the implementation of structuring elements in the cooperative learning environment: a) cooperative learning with no additional metacognitive structuring elements and b) cooperative learning with additional metacognitive structuring elements. 258 seventh-grade lower secondary school students took part in the study. Data were collected in a quasi-experimental pretest-posttest design with a test on socioscientific decision making. Data analyses were conducted with the Rasch Partial Credit Model. As a first step, we analysed whether it was possible to use the same test instrument across different points of measurement without restrictions in terms of validity. The presented analyses display an exemplary procedure to deal with pretest-posttest designs on the basis of IRT modelling. With respect to the increase in decision making competence significant gains could be found. However, differential effects for the treatment condition with additional metacognitive elements could not be found.

Keywords: socioscientific decision making, cooperative learning, metacognitive training, measurement of change

1 Einleitung

Bewertungskompetenz hat seit der Verabschiedung der Bildungsstandards für den mittleren Schulabschluss (KMK, 2004) auf nationaler und seit in Kraft treten neuer Ländercurricula auch auf Länderebene (vgl. u.a. Niedersächsisches Kultusministerium, 2007) eine zentrale Bedeutung für den Biologieunterricht erhalten. Bewertungskompetenz lässt sich in Anlehnung an die Definition von Kompetenzen als kognitive, kontextspezifische Leistungsdispositionen (vgl. auch Hartig & Klieme, 2006; Koepen, Hartig, Klieme & Leutner, 2008) als die Fähigkeit beschreiben, sich in komplexen Problem- und Entscheidungssituationen im Kontext Nachhaltiger Entwicklung bei mehreren gleichwertigen Handlungsoptionen systematisch und begründet entscheiden zu können (vgl. u.a. Bayrhuber et al., 2007). Typischerweise sind derartige komplexe Problem- und Entscheidungssituationen durch die Abwesenheit *einer* optimalen Handlungsoption gekennzeichnet. Vielmehr müssen mehrere gleichwertige, teilweise in Konflikt zueinander stehende Handlungsoptionen unter Berücksichtigung mehrerer Interessengruppen und damit Perspektiven miteinander verglichen bzw. gegeneinander abgewogen werden (Eggert & Bögeholz, 2006). Eine erfolgreiche Bearbeitung von komplexen Problem- und Entscheidungssituationen erfordert einen naturwissenschaftlichen Unterricht, der Lernprozesse auf die Interaktion von Schülerinnen und Schülern ausrichtet und Diskussions- und Aushandlungsprozesse fördert. Derartige kommunikative Aushandlungsprozesse können durch kooperative Lernarrangements gefördert werden (vgl. u.a. Johnson, Johnson & Stanne, 2000; Slavin, 1980; Sumfleth, Wild, Rumann & Exeler, 2002; Sumfleth, Nicolai & Rumann, 2004).

Zahlreiche empirische Studien zum kooperativen Lernen konnten den positiven Einfluss derartiger Lernarrangements auf beispielsweise Motivation, Leistungsbereitschaft und Lernerfolg nachweisen (ebd.). Dennoch

ist Kooperatives Lernen per se nicht zwangsläufig erfolgreicher als individualisierte Lernumgebungen. Ein Gelingen Kooperativen Lernens ist vielmehr von zahlreichen Einflussfaktoren abhängig, wobei dem Grad der Strukturierung von Lernprozessen eine zentrale Rolle zukommt (Sumfleth et al., 2004; Huber, 1995).

Eine Möglichkeit, kooperative Lernumgebungen zu strukturieren, ist der Einsatz von metakognitiven bzw. metastrategischen Trainings (vgl. u.a. Gonzales-Weil, 2006; Kramarski, Mevarech & Arami, 2002). Ziel dabei ist eine Förderung der Reflexionsfähigkeit von Schülerinnen und Schülern über den Lerngegenstand sowie ihre Lernprozesse. Die Fähigkeit, die eigenen sowie die Entscheidungsfindungsprozesse anderer zu reflektieren und dabei eventuelle Fehler zu korrigieren, ist für Bewertungskompetenz zentral (Eggert & Bögeholz, 2006). Da ein positiver Einfluss von metakognitiven bzw. metastrategischen Trainings auf den Lernerfolg in einigen Studien nachgewiesen werden konnte (Mevarech & Kramarski, 2003; Kramarski et al., 2002), kann erwartet werden, dass sich ein derartiges Training auch positiv auf den Kompetenzerwerb im Hinblick auf Bewertungskompetenz auswirkt.

2 Theoretische Grundlagen

2.1 Kooperatives Lernen im Unterricht

Kooperatives Lernen hat besonders im internationalen Diskurs eine jahrzehntelange Tradition (vgl. u.a. Cohen, 1994; Johnson et al., 2000; Slavin, 1980). Im nationalen Kontext sind v.a. die Arbeiten von Renkl und Mandl bzw. Renkl (1995, 1997) und Krause (2007) zu nennen. Untersuchungen zu kooperativen Lernumgebungen speziell im naturwissenschaftlichen Unterricht beziehen sich dabei v.a. auf den Chemie- (vgl. u.a. Sumfleth et al., 2004; Sumfleth et al., 2002; Eilks, Witteck, Rumann & Sumfleth, 2005; Walpuski & Sumfleth, 2007; Tepner, Roeder, & Melle, 2009) aber auch in einigen Studien

auf den Physik- (vgl. u.a. Berger & Hänze, 2004; Berger, 2007) und Biologieunterricht (vgl. Gonzales-Weil & Harms, 2003; Gonzales-Weil, 2006).

Bei allen Unterrichtsmethoden, die unter das Prinzip des Kooperativen Lernens subsummiert werden, zeichnen sich Lernprozesse durch gegenseitige Abhängigkeiten und daraus resultierende individuelle Verantwortungsübernahme aus. Somit hat kooperatives Lernen, neben der Förderung von Lernerfolg, v.a. auch das Ziel soziale Kompetenzen zu fördern (Gillies, 2004). Zahlreiche empirische Untersuchungen konnten zeigen, dass Kooperatives Lernen einen positiven Einfluss auf Motivation für den Lerngegenstand, Lernerfolg sowie soziale Kompetenzen haben kann (Johnson et al., 2000; Slavin, 1980; Cohen, 1994; Sumfleth et al., 2004; Gillies, 2004). Darüber hinaus konnte gezeigt werden, dass ein auf den Basisprinzipien Kooperativen Lernens gestalteter Unterricht (vgl. u.a. Green & Green, 2005) einem, beispielsweise durch Frontalunterricht gekennzeichneten Unterricht überlegen ist (vgl. u.a. Sumfleth et al., 2004).

Oftmals wurden diese positiven Ergebnisse dazu herangezogen, den Einsatz Kooperativen Lernens als Unterrichtsmethode generell anderen Unterrichtsmethoden vorzuziehen (vgl. u.a. Berger & Hänze, 2004). Renkl und Mandl sprechen in diesem Zusammenhang auch von einer „euphorischen Propagierung kooperativen Lernens“ (Renkl & Mandl, 1995, S. 293). Meta-Analysen bzw. Überblicksartikel zu bestehender Forschung zu Kooperativen Lernen kommen in der Regel auch zu den erwarteten positiven Ergebnissen, wenn auch mit unterschiedlich großen Effekten (vgl. u.a. Johnson et al., 2000; Gräsel & Gruber, 2000). Dennoch gibt es auch Studien, die keinen oder nur einen geringen Vorteil Kooperativen Lernens gegenüber beispielsweise lehrerzentrierter Instruktion (vgl. u.a. Gräsel, 1997; Leutner, 1992; Weinert, 1997; Eilks, 2003) oder aber individualisierten Lernumgebungen mit zusätzlichen Feedbackmaßnahmen zeigten (Krause, 2007). So zeigten Berger und Hän-

ze für den Physikunterricht, dass der Einsatz des Gruppenpuzzles im Rahmen einer kooperativen Lernumgebung nicht zu besseren Lernleistungen führte. Lediglich bei den Themen, die während der Gruppenpuzzles in der Expertengruppe bearbeitet wurden, war die Lernleistung gegenüber der Kontrollgruppe mit Frontalunterricht überlegen (Berger & Hänze, 2004). Des Weiteren konnte für den Chemieunterricht gezeigt werden, dass Kooperatives Lernen im Gegensatz zu klassischem Frontalunterricht zwar einen positiveren Einfluss auf die Fachleistung hat. Dennoch konnten zusätzliche Videoanalysen zeigen, dass dies nicht zwangsläufig an qualitativ besseren Gruppenarbeitsphasen liegen muss (Sumfleth et al., 2004). Damit Kooperatives Lernen erfolgreich sein kann, müssen vielmehr einige zentrale Voraussetzungen gegeben sein (vgl. u.a. Renkl & Mandl, 1995; Slavin, 1980; Walpuski & Sumfleth, 2007). Huber (2001) reduziert diese Voraussetzungen auf drei Hauptfaktoren: Spielraum für Entscheidungen / Autonomieerleben, wechselseitige Verantwortung in der Gruppe durch positive Interdependenz sowie individuelle Verantwortlichkeit für die Gruppenleistung. Walpuski und Sumfleth (2007) diskutieren, auf Basis der Arbeiten von Renkl (1995), sechs Bedingungen erfolgreicher Gruppenarbeit. Das Element der Strukturierung wird dabei von vielen Autorinnen und Autoren für die Konzeption von kooperativen Unterrichtsskripts sowie einer erfolgreichen Implementation als zentral herausgestellt (vgl. ebd.; Sumfleth et al., 2004). Kooperative Lernformen sind typischerweise durch ein hohes Maß an Schülerautonomie gekennzeichnet. Ist der Grad der Autonomie jedoch zu hoch, dann können Gruppenprozesse auch ineffizient und oberflächlich werden. Eine Balance zwischen Offenheit und Strukturiertheit ist somit eine Voraussetzung für ein erfolgreiches Gelingen kooperativer Lernprozesse (ebd.).

2.2 Strukturierungshilfen in kooperativen Lernumgebungen

Ein Beispiel für eine zusätzliche Strukturierung kooperativer Lernarrangements ist der Einsatz von metakognitiven Trainings. Der Einfluss derartiger Trainings auf den Lernerfolg von Schülerinnen und Schülern wurde bis jetzt v.a. für den Bereich des Lesens (Hasselhorn & Körkel, 1983, 1986; Cross & Paris, 1988), des Problemlösens (King, 1999) oder des Mathematikunterrichts beschrieben (Schoenfeld, 1987, 1992). Besonders für den Bereich des Mathematikunterrichts wurden derartige Trainings in kooperative Kleingruppenarbeit integriert (vgl. Schoenfeld, 1992; Mevarech & Kramarski, 2003; Kramarski et al., 2002). Ziel dieser kooperativ-metakognitiven Lernarrangements ist dabei, die Interaktion von Schülerinnen und Schülern im Hinblick auf die Elaboriertheit von Erklärungen zu verbessern. Kramarski und Mevarech (2002) konnten für mathematisches Problemlösen zeigen, dass ein kooperativ-metakognitives Lernarrangement im Vergleich mit kooperativen oder individualisierten Lernarrangements den größten Effekt in Bezug auf Lernerfolg hatte. Durch

die metakognitive Instruktion wurden die Schülerinnen und Schüler dazu angeregt, über ihren mathematischen Problemlöseprozess zu reflektieren und dabei die Aufgabe in ihrer Gesamtheit zu durchdenken. Konkret kamen dabei metakognitive Fragen zum Einsatz, die die Schülerinnen und Schüler zu Erklärungen und zum Elaborieren ihrer Denkprozesse anregen sollten. Der positive Einfluss von Fragetechniken, die metakognitive Prozesse anstoßen, wurde bereits mehrfach nachgewiesen (vgl. u.a. King, 1999; für einen Überblick siehe Rosenshine, Meister & Chapman, 1996). So setzte King in der Methode des *Guided Peer Questioning* Verständnis- („comprehension“) und Verknüpfungsfragen („connection questions“) ein. Durch derartige Fragen, die auch als „thought provoking“ bezeichnet werden, soll das Nachdenken über den Lerngegenstand angeregt werden (King, 1994, zitiert in Souvignier, 2005). In einer Studie zum mathematischen Problemlösen wurden diese Fragetypen um Strategie- und Reflexionsfragen erweitert (Kramarski et al., 2002). Die Funktion sowie Beispiele für die einzelnen Fragetypen sind in Tabelle 1 zusammengestellt.

Tab. 1: Übersicht der Fragetypen aus der Studie von Kramarski et al. (2002)

Fragentyp	Funktion	Beispiel
Verständnisfragen	Reflexion über die Aufgabe bevor sie gelöst wird	• Um welches Problem geht es in der Aufgabe?
		• Welche Frage soll beantwortet werden?
Strategiefragen	Reflexion über die für die Aufgabe notwendigen kognitiven Strategien	• Welche Strategie kann zur Lösung dieses Problems genutzt werden?
		• Wie kann ich die Informationen verarbeiten, um das Problem zu lösen?
Verknüpfungsfragen	Reflexion über Gemeinsamkeiten und Unterschiede von bearbeiteten Aufgaben	• Inwiefern ähnelt das Problem anderen Problemen die Du bereits gelöst hast?
Reflexionsfragen	Reflexion während des Lernprozesses	• Welche Schwierigkeiten treten bei der Bearbeitung des Problems auf?
		• Woran arbeite ich gerade?

3 Fragestellung

In der vorliegenden Interventionsstudie wird der Einfluss zweier kooperativer Treatmentvarianten auf den Lernzuwachs von Schülerinnen und Schülern im Hinblick auf Bewertungskompetenz analysiert. Da kooperative Lernumgebungen auf die Interaktion von Schülerinnen und Schülern und den Austausch und die Reflexion über den Lerngegenstand ausgerichtet sind (s.o.), ist zu vermuten, dass sich beide Treatmentvarianten positiv auf den Kompetenzerwerb der Schülerinnen und Schüler auswirken. Bisherige Studien zu Kooperativem Lernen konnten – mit einigen Ausnahmen (siehe Berger & Hänze, 2004) – in der Regel auch einen derartigen positiven Effekt nachweisen (vgl. u.a. Johnson et al., 2000; Sumfleth et al., 2004). Interventionsstudien zur Förderung mathematischen Problemlösens in kooperativen Lernumgebungen (vgl. Mevarech & Kramarski 2003; Kramarski et al., 2002) haben gezeigt, dass v.a. zusätzliche Strukturierungsmaßnahmen durch metakognitive Fragetechniken einen positiven Einfluss auf den Lernerfolg haben (ebd.). Dabei wird v.a. die Förderung der Reflexionsfähigkeit über den Lerngegenstand bzw. über Lernprozesse hervorgehoben. Ziel des Einsatzes eines metakognitiven Trainings in der vorliegenden Studie ist es, die Reflexionsfähigkeit der Schülerinnen und Schüler im Hinblick auf Bewertungskompetenz, dabei v.a. die Reflexion über Entscheidungsprozesse, zu fördern. Da die Fähigkeit zur Reflexion über Entscheidungsprozesse einen zentralen Aspekt in Bezug auf eine hohe Bewertungskompetenz darstellt, wird angenommen, dass das Treatment mit integriertem metakognitiven Training einer normalen kooperativen Lernumgebung überlegen ist. Konkret liegt der Studie im Hinblick auf die Intervention die folgende zentrale Forschungsfrage zugrunde:

- Inwiefern hat eine Treatmentvariante Kooperativen Lernens mit zusätzlichem metakognitiven Training einen stärkeren

positiven Einfluss auf die Bewertungskompetenz der Schülerinnen und Schüler als eine Treatmentvariante Kooperativen Lernens ohne eine zusätzliche Strukturierungshilfe?

Neben dieser inhaltlich fokussierten Forschungsfrage hat die Studie die Bearbeitung einer weiteren, methodischen Frage zum Ziel. Da in der vorliegenden Studie Bewertungskompetenz im Rahmen eines Prä-Posttest-Designs mit demselben Messinstrument erhoben wird, stellt sich zunächst die Frage nach der Vergleichbarkeit der Messungen zu den beiden Messzeitpunkten. Konkret lautet die methodische Frage:

- Inwiefern kann die Qualität der Bewertungskompetenz der Schülerinnen und Schüler zum zweiten Messzeitpunkt als ein Lernzuwachs im Vergleich zum ersten Messzeitpunkt interpretiert werden?

Damit wird eine für Veränderungsmessungen typische Fragestellung – die der Voraussetzung der Messinvarianz – angesprochen (vgl. u.a. Rost, 2004). Erst wenn zumindest eine partielle Messinvarianz der Itemparameter, und damit eine Vergleichbarkeit beider Messzeitpunkte, sichergestellt ist, kann die differentielle Wirksamkeit der beiden Treatments untersucht werden.

4 Methodik

4.1 Stichprobe und Design

Die hier vorgestellte Interventionsstudie wurde im Rahmen des nationalen Projekts Biologie im Kontext an sieben Gymnasien in Süd-Ost Niedersachsen im Rahmen eines Prätest-Posttest-Kontrollgruppendedesigns durchgeführt. An der Untersuchung nahmen insgesamt 258 Schülerinnen und Schüler der siebten Jahrgangsstufen teil (mittleres Alter 12,84; davon 128 Mädchen). Die beiden Treatmentvarianten wurden von je fünf Lehrkräften unterrichtet. Die Lehrkräf-

Tab. 2: Inhalte und geförderte Kompetenzen der Unterrichtseinheit zum Thema Fließgewässerbewertung für beide Treatmentvarianten

Stunde	Inhalt	schwerpunktmäßig geförderte Kompetenz	Einsatz von Strukturierungshilfen bei KOOP META
1	Kaufentscheidung Lebensmittel (Apfelsorten)	Bewertung	Verständnis- und Strategiefragen
2	KOOP: Vertiefung Kaufentscheidung Lebensmittel KOOP META: Umgang mit Gruppenlogbuch am Beispiel Kaufentscheidung Lebensmittel	Bewertung	Verständnis-, Strategie- und Reflexionsfragen, Einführung in den Umgang mit Gruppenlogbuch
3	Kriterien zur Bestimmung des ökologischen Zustands eines Fließgewässers	Erkenntnisgewinnung	---
4	Flussverläufe von der Quelle bis zur Mündung	Fachwissen	---
5	Tiere im und am Bach	Fachwissen	---
6	Nahrungsketten und -netze	Fachwissen	---
7	Stoffkreisläufe und Nahrungsbeziehungen im Bach	Fachwissen	---
8	Anthropogene Beeinflussung	Fachwissen, Bewertung	Verständnisfragen, Einsatz Logbuch
9	Vorbereitung und Planung einer Exkursion	Fachwissen, Bewertung	Strategiefragen, Einsatz Logbuch
	<i>Exkursion</i>	<i>Fachwissen, Erkenntnisgewinnung</i>	
10	Auswertung Exkursion	Fachwissen, Bewertung	
11	Auswertung Exkursion	Fachwissen, Bewertung	
12	Maßnahmen zur Fließgewässergestaltung	Bewertung	Strategie-, Verknüpfungs- und Reflexionsfragen, Einsatz Logbuch
13	Maßnahmen zur Fließgewässergestaltung	Bewertung	Strategie-, Verknüpfungs-, Reflexionsfragen, Einsatz Logbuch

te wurden zufällig den beiden Treatmentvarianten zugeordnet. Keine der Lehrkräfte hatte Vorerfahrungen im Bereich zusätzlicher Strukturierungsmaßnahmen in kooperativen Lernumgebungen oder im Einsatz von beispielsweise Lerntagebüchern. Insgesamt waren 118 Schülerinnen und Schüler der Treatmentvariante Kooperatives Lernen und 140 Schülerinnen und Schüler der Treatmentvariante Kooperatives Lernen mit zusätzlichem metakognitiven Training zugeordnet.

4.2 Anlage der Untersuchung

In intensiver Zusammenarbeit zwischen Lehrkräften und Fachdidaktikerinnen wurde eine Unterrichtseinheit zum Thema Fließgewässerbewertung für die siebte Jahrgangsstufe entwickelt. Die Unterrichtseinheit beinhaltete dabei sowohl Stunden zum ökologischen Basiswissen, als auch mehrere Stunden zur Vermittlung und Förderung von Bewertungskompetenz an unterschiedlichen Kontexten. Die Stunden zur Bewertungskompetenz bauten dabei kumulativ aufeinander auf. An einem relativ einfachen Beispiel (Kaufentscheidung Lebensmittel) wurden zentrale Aspekte von Bewertungskompetenz erarbeitet. Diese wurden an einem weiteren Beispiel geübt (ökologische Bewertung von Fließgewässern) und letztendlich auf ein drittes Beispiel (Fließgewässergestaltung) transferiert (vgl. Tab. 2). Die Unterrichtseinheit wurde methodisch in zwei Treatmentvarianten zum Kooperativen Lernen realisiert: zum einen Kooperatives Lernen ohne zusätzliche Strukturierungshilfe (KOOP), zum anderen Kooperatives Lernen mit zusätzlichem metakognitiven Training (KOOP META). Die Dauer der Unterrichtseinheit war für beide Treatmentvarianten gleich. In beiden Treatmentvarianten wurden die gleichen fachlichen Inhalte unterrichtet. In der Treatmentvariante KOOP META wurde ein zusätzliches metakognitives Training integriert. Dementsprechend hatte die Treatmentvariante KOOP innerhalb

der einzelnen Unterrichtsstunden mehr Zeit zur Erarbeitung der fachlichen Inhalte. Einen Überblick über die Inhalte der einzelnen Unterrichtsstunden sowie den Einsatz des metakognitiven Trainings gibt Tabelle 2. Als konkrete Unterrichtsmethoden kamen in beiden Varianten das Gruppenpuzzle sowie das generelle Prinzip des „Think, Pair, Share“ (vgl. u.a. Brüning & Saum, 2008) zum Einsatz.

4.3 Das Treatment Kooperatives Lernen mit metakognitiven Training (KOOP META)

Das zusätzliche metakognitive Training soll Schülerinnen und Schüler in der Reflexion über ihr (strategisches) Wissen und ihre Lernprozesse unterstützen. Dazu kamen Gruppenlogbücher zum Einsatz, die Schülerinnen und Schüler an zentralen Stellen der Unterrichtseinheit in ihrer Reflexion über die zu erwerbenden Kompetenzen unterstützen sollten. Die Gruppenlogbücher wurden ausschließlich im Unterricht ausgefüllt. Die in den Gruppenlogbüchern zu bearbeitenden Fragen wurden auf Basis der Forschungsarbeiten zum mathematischen Problemlösen entwickelt (Kramarski et al., 2002). So wurden die oben beschriebenen vier Fragetypen aufgegriffen und thematisch sowie im Hinblick auf Bewertungskompetenz adaptiert. Außerdem wurden die Reflexionsfragen nicht nur während sondern zusätzlich auch im Anschluss an Lernphasen gestellt. Wenn möglich wurden aus didaktisch-methodischen Gründen sogenannte „W-Fragen“ durch Operatoren ersetzt. In Tabelle 3 sind Beispielfragen bzw. -impulse zu den vier Fragetypen dargestellt.

4.4 Erhebungsinstrument

Der Einfluss der beiden Treatmentvarianten auf den Lernzuwachs der Schülerinnen und Schüler wurde mit Hilfe eines Messinstruments zur Bewertungskompetenz im Rah-

Tab. 3: Beispielfragen zu den vier Fragetypen aus den Gruppenlogbüchern der Treatmentvariante KOOP META

Fragentyp	Beispielfragen
Verständnisfragen / Impulse	<p>Bevor Du mit der Aufgabe beginnst, beantworte die folgenden Frage: <i>Was ist das Ziel der Aufgabe?</i></p> <p>Untersuchungen von Fließgewässern werden durchgeführt, um herauszufinden, ob der Zustand eines Gewässers gut ist oder verbessert werden muss. D.h. man bewertet den Zustand eines Gewässers. <i>Welche Aspekte sind bei der Bewertung eines Fließgewässers zu berücksichtigen?</i></p>
Strategiefragen / Impulse	<p><i>Erstellt in Eurer Gruppe ein Ablaufschema, wie ihr bei einer Bewertung eines Baches / Flusses vorgehen würdet!</i></p> <p><i>Inwiefern hätte man die Aufgabe auch ohne die vorgestellte Methode (der Entscheidungsfindung) bzw. mit einer anderen Methode besser lösen können?</i></p>
Verknüpfungsfragen / Impulse	<p>In den letzten Stunden zum Fließgewässer habt ihr viel über die Struktur und die Lebewesen dieses Ökosystems erfahren. In diesen drei Stunden innerhalb der Einheit ging es auch darum, Entscheidungen zu treffen.</p> <p><i>Wie seid ihr vorgegangen, um zu einer Entscheidung zu gelangen? Notiert die einzelnen Schritte!</i></p> <p>Diese drei Stunden weisen nicht nur Gemeinsamkeiten, sondern auch Unterschiede auf.</p> <p><i>Benennt die Unterschiede, indem ihr die entsprechenden Merkmale in der folgenden Tabelle ankreuzt [Tabelle nicht dargestellt].</i></p>
Reflexionsfragen / Impulse	<p><i>Beschreibt eventuelle Schwierigkeiten, die ihr während der Bearbeitung der Aufgabe habt bzw. hattet!</i></p> <p><i>Welches Wissen benötigt Ihr noch, um die Aufgabe lösen zu können?</i></p> <p><i>Was war das Ziel der Stunde?</i></p>

men einer Prä-Posttestmessung erhoben. Die Prä- und Posttestmessungen fanden jeweils im Biologieunterricht vor bzw. nach der Intervention statt. Es lagen zu beiden Zeitpunkten mindestens 24 Stunden zwischen den Messungen und der Intervention. Das Messinstrument wurde zuvor in mehreren Studien entwickelt (vgl. Eggert, 2008; Eggert & Bögeholz, 2010). Das Messinstrument besteht aus zwei Abschnitten. Zunächst bearbeiten die Schülerinnen und Schüler zwei Problem- und Entscheidungssituationen aus dem Bereich Nachhaltige Entwicklung („Bewertungsaufgaben“). Konkret geht es dabei um die Bewertung

von Handlungsoptionen zur Sicherung des Dorschbestandes in der Ostsee bzw. zur Bekämpfung einer Neophyte an Fließgewässern. Beide Situationen sind dabei typisch für Problem- und Entscheidungssituationen Nachhaltiger Entwicklung: sie weisen keine eindeutig richtige, sondern vielmehr mehrere gleichwertige Handlungsoptionen auf, die alle durch Vor- aber auch Nachteile gekennzeichnet sind. Darüber hinaus zeichnen sich alle Handlungsoptionen durch ein gewisses Maß an Unsicherheit im Hinblick auf Umsetzbarkeit und zukünftigen Erfolg aus. Die Schülerinnen und Schüler werden bei diesen beiden Situationen aufgefordert,

alle Handlungsoptionen miteinander zu vergleichen. Um sich anschließend für eine Option bzw. eine Kombination von Optionen zu entscheiden, sollte dabei eine Entscheidungsstrategie angewendet werden.

Im zweiten Abschnitt werden den Schülerinnen und Schülern Entscheidungsprozesse und Entscheidungen anderer Schülerinnen und Schüler präsentiert. Bei diesem Aufgabenkomplex handelt es sich thematisch um eine Kaufentscheidung für konventionell oder fair gehandelte Schokolade. Es werden dabei unterschiedliche Entscheidungsstrategien präsentiert, die von den Schülerinnen und Schülern beschrieben, verglichen und beurteilt werden sollen („Reflexionsaufgabe“).

In beiden Testabschnitten kommen ausschließlich offene Aufgabenformate zum Einsatz, d.h. die Bewertungskompetenz der Schülerinnen und Schüler wird anhand der Qualität ihrer Antworten erfasst und nicht anhand der Anzahl gelöster Aufgaben. Die Schülerantworten werden insgesamt im Hinblick auf 13 dichotome und polytome Auswertungssitems analysiert. Diese 13 Items teilen sich dabei auf die eigenen Entscheidungen der Schülerinnen und Schüler (Items 1-10), bzw. auf die Beschreibung und Beurteilung der präsentierten Entscheidungen bzw. Entscheidungsstrategien (Items 11-13) auf. Einen Überblick dieser 13 Items gibt Tabelle 4.

4.5 Auswertungsmethodik

In einem ersten Schritt wurden beide Messzeitpunkte getrennt mit Hilfe des Rasch Partial Credit Modells analysiert (Raschmodell für polytome Daten; Masters, 1982). Dabei konnte das Rasch Partial Credit Modell im Vergleich mit Mischverteilungsmodellen zu beiden Zeitpunkten als das am besten auf die Daten passende identifiziert werden. Es bildet daher die Grundlage aller weiteren Analysen. Ein Vergleich der Personenparameter der Schülerinnen und Schüler zu beiden Messzeitpunkten setzt wenigstens die

partielle Invarianz der Itemparameter zu T1 und T2 voraus. Die Untersuchung der Invarianz der Itemparameter (Messinvarianz) erfolgte in zwei Schritten: Erstens wurden die Daten der beiden Messzeitpunkte getrennt voneinander modelliert und die Itemparameter der 13 Items frei geschätzt. Die Modelle wurden jeweils über die Personenparameter identifiziert (Analyseoption *constraints=cases* in Conquest; Wu et al., 1998). Zweitens wurden die Itemparameter zu beiden Messzeitpunkten mit Hilfe eines grafischen Modelltests miteinander verglichen. Ziel dieses Vergleichs war es, Items zu identifizieren, die sich in ihrer Schwierigkeit zu T2 deutlich, d.h. unverhältnismäßig, von der Schwierigkeit zu T1 unterscheiden. Mit derartigen *differential item functioning (dif)* Analysen lassen sich für Veränderungsmessung problematische Items identifizieren (vgl. u.a. Wilson, 2005; Bond & Fox, 2001; Yilmaz, Boone & Anderson, 2004).

Es wurden Warm's *weighted likelihood*-Schätzer (WLE; vgl. Warm, 1989) als Schätzer für die Personenfähigkeit verwendet (vgl. Rost, 2004). Die Schätzungen wurden in zwei getrennten Skalierungsläufen vorgenommen, wobei zunächst die Leistungen zu T1 und im Anschluss hieran zu T2 geschätzt wurden. Entsprechend wurden zur Herstellung einer gemeinsamen Metrik die Itemparameter aus der Schätzung zu T1 in die Skalierung von T2 eingesetzt. Dabei wurden die im Hinblick auf *dif* problematischen Items frei geschätzt (partielle Invarianz). Dieses Vorgehen bedeutet, dass diese Items nicht in die Veränderungsmessung und somit nicht in die Analysen von Leistungszuwächsen eingehen. Sie tragen aber zu einer effizienteren Schätzung der Personenparameter bei. Durch die fixierten Ankerparameter war das Modell zu T2 identifiziert, weshalb auf eine weitere Identifizierung des Modells über Personen- oder Itemparameter verzichtet werden konnte (Analyseoption *constraints=none* in Conquest).

Um die Personenparameter des zweiten Messzeitpunktes mit den Personenparametern des ersten Messzeitpunktes vergleichen zu können, wurden diese am Mittelwert und der

Tab. 4: Beschreibung der zum Scoren von Schülerantworten verwendeten Auswertungsimens

	Item Nr.	Beschreibung	Schülerantworten werden gescort im Hinblick auf ...	Antwort-kategorie	Itemfit Prätest (MNSQ, pro Kategorie)	Itemfit Posttest (MNSQ, pro Kategorie)
Bewertungsaufgabe „Dorsch“	1	Option „Vergrößerte Maschen in Fischnetzen“	Pro- und Kontra-Argumente (0 = keine Argumente, 1 = nur Pro- oder nur Kontra-Argumente, 2 = Pro- und Kontra-Argumente)	0 1 2	0,87 0,94 0,95	0,91 0,93 0,84
	2	Option „Fischfang für bestimmte Zeit komplett einstellen“	Pro- und Kontra-Argumente (siehe Item 1)	0 1 2	0,85 0,95 0,98	0,96 0,92 0,75
	3	Option „Schutzzeiten und Schutzzonen“	Pro- und Kontra-Argumente (s.o.)	0 1 2	1,08 1,06 1,21	1,15 1,06 1,20
	4	Option „Fangquote“	Pro- und Kontra-Argumente (s.o.)	0 1 2	0,82 0,93 0,99	0,80 0,91 0,78
	5	Gewichten	Gewichten von Entscheidungskriterien (0 = Kriterien nicht gewichtet, 1 = Kriterien gewichtet)	1	1,18	1,11
Bewertungsaufgabe „Neophyte“	6	Option „Einsatz eines Pflanzenschutzmittels“	Pro- und Kontra-Argumente (s.o.)	0 1 2	0,87 0,94 0,92	0,89 0,92 0,89
	7	Option „Anpflanzen von Schilf und Schwarzerle“	Pro- und Kontra-Argumente (s.o.)	0 1 2	1,13 1,05 1,17	1,24 1,00 1,04
	8	Option „Beweidung durch Schafe“	Pro- und Kontra-Argumente (s.o.)	0 1 2	0,91 0,93 0,96	1,01 0,95 0,95
	9	Option „Mahd“	Pro- und Kontra-Argumente (s.o.)	0 1 2	0,88 0,93 0,87	0,88 0,90 0,69
	10	Gewichten	Gewichten von Entscheidungskriterien (siehe Item 5)	1	1,12	1,06
Reflexionsaufgabe „Schokoladenkauf“	11	Beschreibung Entscheidungsstrategie 1	Entscheidungsprozess durch Ausschlussverfahren beschreiben (0 = Prozess wird nicht oder nur teilweise beschrieben, 1 = Prozess wird vollständig beschrieben und/oder kritisch reflektiert)	1	1,10	1,31
	12	Beschreibung Entscheidungsstrategie 2	Entscheidungsprozess durch Abwägen der Pro- und Kontra-Argumente beschreiben (0 = Prozess wird nicht oder nur teilweise beschrieben, 1 = Prozess wird vollständig beschrieben)	1	1,07	1,21
	13	Verbesserungsvorschlag für Entscheidungsstrategie 2	Verbesserungsvorschläge für den Entscheidungsprozess durch Abwägen (0 = kein Vorschlag genannt oder nur inhaltlich argumentiert; 1 = methodischer Vorschlag genannt)	1	1,09	1,06

Tab. 5: Analyse der Logitdifferenzen der Itemdeltas der Messzeitpunkte t1 und t2

Item / -step	Itemdelta (T1)	Itemdelta (T2)	Logitdifferenz
1, step1	-0,390	-0,420	0,030
1, step2	1,260	0,570	0,690
2, step1	-0,170	-0,540	0,370
2, step2	0,630	0,550	0,080
3, step 1	-1,510	-1,710	0,200
3, step 2	0,380	-0,360	0,740
4, step 1	-0,100	-0,090	0,010
4, step 2	1,320	0,750	0,570
5	2,920	1,850	1,070
6, step 1	-0,570	-0,740	0,170
6, step 2	0,970	0,520	0,450
7, step 1	-1,410	-1,380	0,030
7, step 2	1,150	0,400	0,750
8, step1	-0,120	-0,520	0,400
8, step2	0,820	-0,040	0,860
9, step 1	-0,150	-0,460	0,310
9, step 2	1,990	0,840	1,150
10	3,880	1,990	1,890
11	3,270	3,300	0,030
12	4,380	4,230	0,150
13	3,750	4,610	0,860

Hinweis: Nach Paek (2002) sind Logitdifferenzen über .638 als hoch zu klassifizieren. Die entsprechenden Logitdifferenzen sind grau markiert (Paek, 2002; zitiert in Wilson, 2005).

Standardabweichung der Personenparameter des ersten Messzeitpunktes standardisiert. Zusätzlich wurden sie auf eine gemeinsame Metrik mit einem Mittelwert von 100 und einer Standardabweichung von 30 gebracht. Für den Treatmentvergleich wurden die Personenfähigkeiten beider Messzeitpunkte im Weiteren klassisch statistisch ausgewertet.

ihrer Itemsteps im Hinblick auf Logitdifferenzen wird dieser grafische Modelltest bestätigt (vgl. Tab. 5). Die Logitdifferenzen von insgesamt acht Items bzw. Itemsteps sind als

5 Ergebnisse

5.1 Messung derselben Kompetenz zu beiden Messzeitpunkten

Im Rahmen der *dif* Analysen wurden die Itemdeltas beider Messzeitpunkte in einem Streudiagramm gegeneinander aufgetragen (Abb. 1). Diejenigen Items, die nicht innerhalb des 95% Konfidenzintervalls liegen, sind im Hinblick auf *dif* problematisch. Somit weißt das dargestellte Streudiagramm mehrere Items als problematisch aus. Analysiert man die Itemdeltas dieser Items bzw.

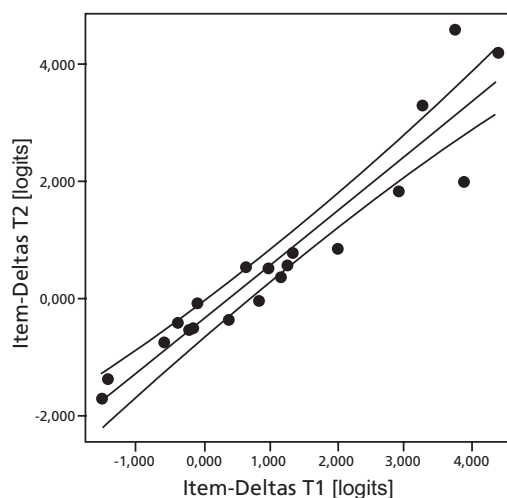


Abb. 1: Untersuchung der eingesetzten Items im Hinblick auf *differential item functioning* mit Hilfe eines grafischen Modelltests.

hoch zu klassifizieren (vgl. Paek, 2002, zitiert in Wilson, 2005).

Bei den Items 1, 3, 7, 8 und 9 handelt es sich um Items, bei denen die Schülerinnen und Schüler Pro- und Kontra- Argumente für die in den Aufgaben präsentierten Handlungsoptionen beschreiben sollen (vgl. auch Tab. 4). Bei diesen Items wird jeweils der zweite Itemstep im Posttest sehr viel leichter, d.h. es wird für Schülerinnen und Schüler relativ einfach, einen Score 2 auf diesen Items zu erhalten (vgl. Tab. 5).

Die weiteren problematischen Items beziehen sich auf das Gewichten von Entscheidungskriterien in einem Bewertungsprozess (vgl. Tab. 4). Auch hier werden diese Items im Posttest unverhältnismäßig einfach. Das Item 13 schließlich bezieht sich auf die Reflexionsaufgabe „Schokoladenkauf“ (vgl. ebd.). Im Gegensatz zu den anderen Items wird dieses Item zum zweiten Messzeitpunkt unverhältnismäßig schwer. Als Konsequenz für die Veränderungsmessung ergibt sich daraus, diese Items bzw. einzelne Itemsteps weiterhin frei zu schätzen.

Nach der oben dargestellten Prüfung der Messinvarianz wurde für den Prätest eine WLE Person Separation Reliabilität von .73 und Cronbach's α von .73 ermittelt. Für den Posttest lag die WLE Person Separation Reliabilität bei .85; Cronbach's α .86.

5.2 Treatmentvergleich

Zu Beginn der Unterrichtseinheit unterschieden sich die beiden Treatmentvarianten nicht signifikant im Hinblick auf ihre Bewertungskompetenz ($t(256) = -1.632$, $p = .104$). Die Ergebnisse einer zweifaktoriellen Varianzanalyse mit Messwiederholung (2 Treatmentgruppen x 2 Testzeitpunkte) zeigten einen höchst signifikanten Haupteffekt der Zeit ($F_{(1,256)} = 17.240$, $p < .001$) sowie einen signifikanten Haupteffekt der Treatmentgruppe ($F_{(1,256)} = 6.136$, $p < .05$). Der vorhergesagte Interaktionseffekt zeigte sich zwar tendenziell, konnte jedoch statistisch nicht nachgewiesen werden ($F_{(1,256)} = 1.751$,

$p > .05$; vgl. Abb. 2). Eine zusätzliche Kovarianzanalyse über die Posttestwerte mit den Prätestwerten als Kovariate ergab ebenfalls einen tendenziellen Effekt der Gruppe. Das 5% Signifikanzniveau wurde dabei nur knapp verfehlt ($F_{(1,256)} = 3,598$, $p = .06$).

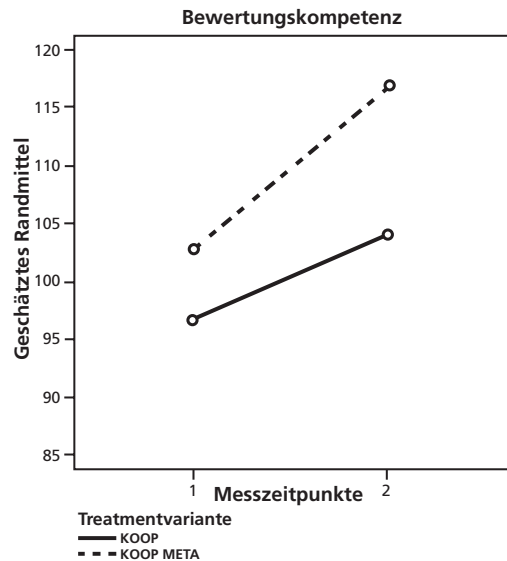


Abb. 2: Leistung im Bezug auf Bewertungskompetenz als Funktion des Messzeitpunktes (für beide Treatmentvarianten).

6 Zusammenfassung und Diskussion

In der vorliegenden Studie wurde die Wirksamkeit zweier kooperativer Lernumgebungen (kooperative Lernumgebung mit bzw. ohne metakognitivem Training) im Hinblick auf die Bewertungskompetenz von Schülerinnen und Schülern im Biologieunterricht untersucht. Die Ergebnisse des Prä-Posttest-Vergleichs zeigen, dass Schülerinnen und Schüler beider Treatmentvarianten die Testaufgaben im Anschluss an die Intervention deutlich erfolgreicher bearbeiten konnten als vorher. Es wurden signifikante Haupteffekte der Zeit (die Kompetenzen nahmen vom Prätest zum Posttest zu) sowie der Gruppenzugehörigkeit (zu beiden Messzeitpunkten war die Gruppe des kombinierten Treatments der Vergleichsgruppe

überlegen) gefunden. Ein Interaktionseffekt – und damit eine differentielle Wirksamkeit – konnte zwar tendenziell beobachtet, jedoch statistisch nicht nachgewiesen werden. Jedoch ist hier anzumerken, dass die Ergebnisse der zusätzlichen Kovarianzanalyse einen signifikanten Effekt der Gruppe nur knapp verfehlten. Das Ausbleiben des vorhergesagten differentiellen Treatmenteffektes mag verschiedene Gründe haben, die z.B. in der Durchführung des Trainings im Schulkontext liegen oder auf Problemen der erforderlichen Messinvarianz der Itemparameter zu den zwei berücksichtigten Messzeitpunkten beruhen.

Ein Vergleich der Itemparameter zu den beiden Messzeitpunkten identifizierte mehrere Items bzw. bei den polytomen Items einzelne Itemsteps als problematisch. Diese Items konnten folglich nicht in die Veränderungsmessung aufgenommen werden. Dies hatte einen relativ großen Informationsverlust zur Folge. Bei den problematischen Items handelt es sich jedoch um Items, die zentrale Aspekte im Hinblick auf eine Förderung von Bewertungskompetenz darstellen. Zum einen ist dies die Fähigkeit sowohl Pro- als auch Kontra- Argumente im Hinblick auf Handlungsoptionen, d.h. Lösungsansätze, für Problemsituationen im Bereich des Umweltschutzes bzw. Nachhaltigen Entwicklung zu beschreiben. Zum anderen gehörte zu diesen Items die Gewichtung von Entscheidungskriterien bei der Bewertung von Handlungsoptionen. Alle genannten Items bzw. Itemsteps wurden zum zweiten Messzeitpunkt für die Schülerinnen und Schüler unverhältnismäßig einfach. Anders verhielt es sich bei den Items für die Reflexionsaufgabe zum Thema „Schokoladenkauf“. Diese Items blieben in etwa gleich schwierig oder wurden zum zweiten Messzeitpunkt schwieriger. Ein Erklärungsansatz hierfür ist, dass sowohl das Bearbeiten von Pro- und Kontra- Argumenten als auch das Gewichten von relevanten Entscheidungskriterien von den Lehrkräften beider Treatmentvarianten besonders stark gefördert wurde. Im Gegensatz dazu kann vermutet werden, dass das

Reflektieren über Bewertungsprozesse vergleichsweise nicht so stark gefördert wurde. Dies muss auch für die Treatmentvariante KOOP META vermutet werden, da durch die varianzanalytischen Untersuchungen keine größere Wirksamkeit nachgewiesen werden konnte. Möglicherweise ist die Treatmentvariante KOOP META dadurch zu optimieren, dass noch stärker und gezielter auf ein Anstoßen der Reflexionsfähigkeit über Bewertungsprozesse geachtet wird.

Um dies umzusetzen, wäre im nächsten Schritt eine Kontrolle der Implementation der Treatments im Biologieunterricht angebracht, etwa durch eine teilnehmende Beobachtung und/oder mit Hilfe von Videoanalysen, was in der vorliegenden Studie nicht realisiert werden konnte. Es fanden aber vor, während und nach der Intervention Treffen mit den beteiligten Lehrkräften statt. Diese Selbstauskünfte der Lehrkräfte brachten hervor, dass es kaum inhaltliche Abweichungen von der Planung gab, jedoch kam der enge Zeitrahmen der geplanten Stunden zur Sprache. Daraus könnte sich ergeben haben, dass die Bearbeitung der Reflexionsaufgaben und insbesondere die Bearbeitung der verschiedenen Fragetypen in der Treatmentvariante KOOP META nicht mit der gewünschten Verarbeitungstiefe stattgefunden hat.

In diesem Zusammenhang muss auch berücksichtigt werden, dass die vorliegende Studie im Rahmen eines Forschungs- und Entwicklungsprojektes durchgeführt wurde. Das Training war somit in einen umfassenden kompetenzorientierten Unterricht eingebettet, indem nicht nur Bewertungskompetenz sondern darüber hinaus auch Fachwissen und Erkenntnisgewinnung gefördert wurde. Somit kamen für die Schülerinnen und Schüler, die an der Intervention teilnahmen, viele neu zu erlernende inhaltliche sowie prozessorientierte Kompetenzen zum Tragen. Dieser Aspekt kann möglicherweise dazu geführt haben, dass das metakognitive Training nicht optimal gewirkt hat. Als Konsequenz hieraus ergibt sich für die Optimierung des Trainings sicherlich

eine Fokussierung auf die im Kern zu fördernden Kompetenzen. Nichtsdestotrotz ist die ökologische Validität des gewählten Ansatzes auch eine hervorzuhebende Stärke. Somit konnte zumindest gezeigt werden, dass beide Interventionsgruppen durch die Intervention in ihrer Bewertungskompetenz gefördert werden konnten. Der tendenzielle Effekt der zusätzlichen metakognitiven Strukturierungshilfen lässt bei den vorgeschlagenen Treatmentoptimierungen (s.o.) vermuten, dass zukünftig differentielle Effekte nachgewiesen werden können.

Danksagungen

Die Interventionsstudie wurde im Rahmen des vom Bundesministerium für Bildung und Forschung (BMBF) geförderten Projekts *Biologie im Kontext* durchgeführt. Ein herzlicher Dank gilt dabei allen beteiligten Lehrkräften sowie Schülerinnen und Schülern.

Die Auswertung der Ergebnisse sowie die Vorbereitung dieser Veröffentlichung wurde ermöglicht durch eine Sachbeihilfe der Deutschen Forschungsgemeinschaft (Kennz.: BO 1730/3-2) im Schwerpunktprogramm „Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen“ (SPP 1293).“

Literatur

- Bayrhuber, H., Bögeholz, S., Elster, D., Hammann, M., Hößle, C., Lücken, M., Mayer, J., et al. (2007). *Biologie im Kontext – Ein Programm zur Kompetenzförderung durch Kontextorientierung im Biologieunterricht und zur Unterstützung von Lehrprofessionalisierung*. *MNU*, 60(5), 282-286.
- Berger, R. & Hänze, M. (2004). Das Gruppenpuzzle im Physikunterricht der Sekundarstufe 2 – Einfluss auf Motivation, Lernen und Leistung. *Zeitschrift für Didaktik der Naturwissenschaften*, 10, 205-219.
- Berger, R. (2007). Das Gruppenpuzzle im Physikunterricht der Sekundarstufe 2 – eine empirische Untersuchung auf Grundlage der Selbstbestimmungstheorie der Motivation. *Studien zum Physik- und Chemielernen* (70). Berlin: Logos.
- Bond, T. G. & Fox, C. M. (2001). *Applying the Rasch Model: Fundamental Measurement in the Human Sciences*. Mahwah NJ: Lawrence Erlbaum Associates.
- Brüning, L. & Saum, T. (2008). *Erfolgreich unterrichten durch kooperatives Lernen. Strategien zur Schüleraktivierung*. Essen: Neue Deutsche Schule Verlagsgesellschaft.
- Cohen, E. G. (1994). Restructuring the classroom: Conditions for productive small groups. *Review of Educational Research*, 64, 1-35.
- Cross, D. R. & Paris, S. C. (1988). 'Developmental and instrumental analysis of children's metacognition and reading comprehension', *Journal of Educational Psychology*, 80, 131-142.
- Eggert, S. (2008). Bewertungskompetenz für den Biologieunterricht – Vom Modell zur empirischen Überprüfung, verfügbar über <http://webdoc.sub.gwdg.de/diss/2008/eggert/eggert.pdf>, (abgerufen am: 22.07.2009).
- Eggert, S. & Bögeholz, S. (2006). Göttinger Modell der Bewertungskompetenz – Teilkompetenz "Bewerten, Entscheiden und Reflektieren" für Gestaltungsaufgaben Nachhaltiger Entwicklung. *Zeitschrift für Didaktik der Naturwissenschaften*, 12, 177-199.
- Eggert, S. & Bögeholz, S. (2010). Students' Use of Decision-Making Strategies With Regard to Socioscientific Issues – An Application of the Rasch Partial Credit Model. *Science Education*, 94, 230-258.
- Eilks, I. (2003). Kooperatives Lernen im Chemieunterricht (Teil1). *MNU*, 56(1), 51-55.
- Eilks, I., Witteck, T., Rumann, S. & Sumfleth, E. (2005). Kooperatives Lernen. *Naturwissenschaften im Unterricht Chemie*, 16, 6-11.

- Gillies, R. (2004). The effects of cooperative learning on junior high school students during small group learning. *Learning and Instruction, 14*, 197-213.
- Gonzales Weil, C. & Harms, U. (2003). Verständnisentwicklung im Bereich Zellbiologie – eine Untersuchung über einen Zusammenhang von Metakognition und Verstehensprozessen. In A. Bauer et al. (Hrsg.), *Entwicklung von Wissen und Kompetenzen*. Internationale Tagung der Sektion Biologiedidaktik im VDBiol 14.-19. September 2003. IPN, Kiel, 159-162.
- Gonzales Weil, C. (2006). *Zusammenhang zwischen Konzeptwechsel und Metakognition: Empirische Untersuchungen über Verstehensprozesse im Bereich Zellbiologie in der 9. Jahrgangsstufe einer chilenischen Oberstufe*. Berlin: Logos Verlag.
- Gräsel, C. (1997). *Problemorientiertes Lernen*. Göttingen: Hogrefe.
- Gräsel, C. & Gruber, H. (2000). Kooperatives Lernen in der Schule. Theoretische Ansätze – Empirische Befunde – Desiderate in der Lehramtsausbildung. In N. Seibert (Ed.), *Unterrichtsmethoden* (pp. 161-176). Opladen: Klinkhardt.
- Green, N. & Green, K. (2005). *Kooperatives Lernen im Klassenraum und im Kollegium: Das Trainingsbuch*. Seelze: Kallmeyersche Verlagsbuchhandlung.
- Hasselhorn, M. & Körkel, J. (1983). Gezielte Förderung der Lernkompetenz am Beispiel der Textverarbeitung. *Unterrichtswissenschaft, 11*, 370-382.
- Hasselhorn, M. & Körkel, J. (1986). Metacognitive versus traditional reading instructions: The mediating role of domain-specific knowledge on children's text processing. *Human Learning, 5*, 75-90.
- Hartig, J. & Klieme, E. (2006). Kompetenz und Kompetenzdiagnostik. In K. Schweizer (Ed.), *Leistung und Leistungsdiagnostik* (pp. 127-143). Heidelberg: Springer Medizin Verlag.
- Huber, G. L. (1995). Lernprozesse in Kleingruppen: Wie kooperieren die Lerner? *Unterrichtswissenschaft, 23*, 316-331.
- Johnson, D. W., Johnson, R. T. & Stanne, M. B. (2000). *Cooperative Learning Methods: A Meta-Analysis*, [verfügbar über: <http://www.co-operation.org/pages/cl-methods.html>, Abrufdatum: 19.07.2009]
- King, A. (1999). Discourse patterns for mediating peer learning. In A. M. Donnell & A. King (Eds.), *Cognitive perspectives on peer learning* (pp. 87-115). Mahwah, NJ: Erlbaum Associates.
- KMK (2004). *Bildungsstandards im Fach Biologie für den Mittleren Schulabschluss* (online), www.kmk.org/schul/Bildungsstandards/Biologie_MSA_16-12-04.pdf (abgerufen am 15. Juni 2004).
- Koepfen, K., Hartig, J., Klieme, E. & Leutner, D. (2008). Current issues in competence modeling and assessment. *Journal of Psychology, 216*(2), 61-73.
- Kramarski, B., Mevarech, Z. R. & Arami, M. (2002). The effects of metacognitive instruction on solving mathematical authentic tasks. *Educational Studies in Mathematics, 49*, 225-250.
- Krause, U. M. (2007). *Feedback und kooperatives Lernen*. Münster: Waxmann.
- Leutner, D. (1992). *Adaptive Lehrsysteme*. Weinheim: Psychologie Verlags Union.
- Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika, 47*, 149-174.
- Mevarech, Z. R. & Kramarski, B. (2003). The effects of metacognitive training versus worked-out examples on students' mathematical reasoning. *British Journal of Educational Psychology, 73*, 449-471.
- Niedersächsisches Kultusministerium (2007). Kerncurriculum für das Gymnasium Schuljahrgänge 5-10. Naturwissenschaften, www.cuvo.nibis.de (abgerufen am 08.11.10).
- Renkl, A. & Mandl, H. (1995). Kooperatives Lernen: Die Frage nach dem Notwendigen und dem Ersetzbaren. *Unterrichtswissenschaft, 23*, 292-300.
- Renkl, A. (1997). *Lernen durch Lehren: Wirkungsmechanismen beim kooperativen Lernen*. Wiesbaden: DUV.
- Rosenshine, B., Meister, C. & Chapman, S. (1996). Teaching students to generate questions: A review of the intervention studies. *Review of Educational Research, 66*, 181-221.
- Rost, J. (2004). *Testtheorie und Testkonstruktion*. Bern, Switzerland: Hans Huber.
- Schoenfeld, A. H. (1987). 'What's all the fuss about metacognition?'. In A. H. Schoenfeld, (Ed.), *Cognitive Science and Mathematics Education* (pp. 19-216). Erlbaum, Hillsdale, NJ.
- Schoenfeld, A. H. (1992). 'Learning to think mathematically: problem solving, metacognition, and sense making in mathematics'. In D. A. Grouws (Ed.), *Handbook of Research on Mathematics Teaching and Learning* (pp. 165-197). MacMillan, New York
- Slavin, R. E. (1980). Cooperative Learning. *Review of Educational Research, 50*, 315-342.
- Souvignier, E. & Kronenberger, J. (2005). Fragen und Erklärungen beim kooperativen Lernen in Grundschulklassen. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 37*(2), 91-100.

- Sumfleth, E., Wild, E., Rumann, S. & Exeler, J. (2002). Wege zur Förderung der naturwissenschaftlichen Grundbildung im Chemieunterricht: kooperatives Problemlösen im schulischen und familialen Kontext zum Themenbereich Säure-Base. *Zeitschrift für Pädagogik*, 45, Beiheft, 207-221.
- Sumfleth, E., Nicolai, N. & Rumann, S. (2004). Schulische und häusliche Kooperation im Chemieanfangsunterricht. In J. Doll & M. Prenzel (Eds.), *Bildungsqualität von Schule: Lehrerprofessionalisierung, Unterrichtsentwicklung und Schülerförderung als Strategien der Qualitätsverbesserung* (pp. 284-302). Münster: Waxmann.
- Tepner, M., Roeder, B. & Melle, I. (2009). Effektivität des Gruppenpuzzles im Chemieunterricht der Sekundarstufe I. *Zeitschrift für Didaktik der Naturwissenschaften*, 15, 7-29.
- Walpuski, M. & Sumfleth, E. (2007). Strukturierungshilfen und Feedback zur Unterstützung experimenteller Kleingruppenarbeit im Chemieunterricht. *Zeitschrift für Didaktik der Naturwissenschaften*, 13, 181-195.
- Warm, T.A. (1989) Weighted Likelihood Estimation of Ability in Item Response Theory. *Psychometrika*, 54, 427-450.
- Weinert, F. E. (1997). Lerntheorien und Instruktionsmodelle. In F. E. Weinert (Hrsg.): *Psychologie des Lernens und der Instruktion* (Bd. 2). Göttingen: Hogrefe, 1-48.
- Wilson, M. (2005). *Construction Measures: An Item Response Modeling Approach*. Mahwah, New Jersey, London. Lawrence Erlbaum Associates.
- Wu, M. L., Adams, R. J., & Wilson, M. R. (1998): ACER ConQuest. Generalised Item Response Modelling Software. Melbourne: Acer Press.
- Yilmaz, Ö., Boone, J. W., & Anderson, O. H. (2004). Views of Elementary and Middle School Turkish Students toward Environmental Issues. *International Journal of Science Education*, 26, 1527-1546.

Kontakt

Dr. Sabina Eggert
Georg-August-Universität Göttingen
Didaktik der Biologie
Waldweg 26
37073 Göttingen
seggert1@gwdg.de

Autoreninformation

Dr. Sabina Eggert ist wissenschaftliche Mitarbeiterin in der Abteilung Didaktik der Biologie an der Georg-August-Universität Göttingen.
Dr. Susanne Bögeholz ist Professorin für Didaktik der Biologie an der Georg-August-Universität Göttingen.

Dr. Rainer Watermann ist Professor für Schulpädagogik und Empirische Schulforschung an der Georg-August-Universität Göttingen.

Dr. Marcus Hasselhorn ist Professor für Psychologie mit dem Schwerpunkt Bildung und Entwicklung am Deutschen Institut für Internationale Pädagogische Forschung (DIPF) in Frankfurt.