

Scaffolding beim Forschenden Lernen

Eine empirische Untersuchung zur Wirkung von Lernunterstützungen

Julia Arnold¹  · Kerstin Kremer¹ · Jürgen Mayer²

Eingegangen: 24. Februar 2016 / Angenommen: 18. November 2016 / Online publiziert: 21. Dezember 2016
© Der/die Autor(en) 2017. Dieser Artikel ist eine Open-Access-Publikation.

Zusammenfassung In der vorliegenden Studie wird untersucht, inwiefern sich die Nutzung von Lernunterstützungen (*Scaffolds*) des prozeduralen Wissens und des Methodenwissens beim Forschenden Lernen eignen, um das Wissenschaftliche Denken, das Methodenwissen und das Fachwissen zu fördern und zudem die kognitive Belastung zu reduzieren. Dazu wurde eine 2×2 -quasi-experimentelle Interventionsstudie mit Pre- und Post-Test in einer 11. Jgst. zum Thema Enzymatik ($N = 220$) durchgeführt. Als unabhängige Variablen wurden theoriebasiert die Unterstützungsformate gestufte Lernhilfen (*Forschertipps*) und diskursiv-reflexive Szenarien (*Concept Cartoons*) gewählt. Als abhängige Variablen wurden das Wissenschaftliche Denken, das Methodenwissen und das Fachwissen sowie die kognitive Belastung der Schüler zu den jeweiligen Testzeitpunkten erhoben. Es zeigt sich, dass sich beide Lernunterstützungen sowie deren Kombination eignen, den Erwerb des Wissenschaftlichen Denkens zu fördern sowie die kognitive Belastung zu reduzieren. Eine Wirksamkeit der Lernunterstützungen in den Bereichen Methodenwissen und Fachwissen konnte nur teils bestätigt werden.

Schlüsselwörter Forschendes Lernen · Interventionsstudie · Wissenschaftliches Denken · Problemlösen · Scaffolding

✉ Julia Arnold
arnold@ipn.uni-kiel.de

¹ Didaktik der Biologie, Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik (IPN), Universität Kiel, Olshausenstraße 62, 24118 Kiel, Deutschland

² Universität Kassel, Heinrich-Plett-Str. 40, 34132 Kassel, Deutschland

Scaffolding in Inquiry Learning

An Empirical Study on the Impact of Learning Support

Abstract The present study examines if and how the use of scaffold-formats to support the procedural knowledge and methodological knowledge (procedural understanding) in inquiry learning can promote scientific thinking, methodological knowledge and content knowledge and also reduce cognitive load. For this purpose a 2×2 quasi-experimental intervention study with pre- and post-test in a 11th grade about enzymology ($N = 220$) was conducted. The scaffold-formats worked examples (*Research Tips*) and discursive reflexive scenarios (*Concept Cartoons*) were chosen as independent variables. According to theory they are suitable to promote procedural and declarative knowledge respectively. As dependent variables scientific thinking, methodological knowledge and content knowledge as well as cognitive load were tested at each test time. It turns out that both scaffold-formats and combinations thereof are suitable to promote the acquisition of scientific thinking and reduce cognitive load. The effectiveness of the scaffolds in the areas of methodological knowledge and content knowledge could be confirmed only partly.

Keywords Inquiry Learning · Intervention study · Inquiry Competence · Problem solving · Scaffolding

Bedeutung des Forschenden Lernens

Wissen ist in unserem digitalen Zeitalter leicht zugänglich, schnell verfügbar und entwickelt sich rapide weiter (AAAS 2009). Daher sollten Lernende einerseits in der Lage sein, vorhandenes Wissen bzw. vorhandene Erkenntnisse kritisch reflektieren zu können und andererseits die Fähig-

keit erwerben, sich im Sinne lebenslangen Lernens Wissen selbst anzueignen und generell das naturwissenschaftliche Vorgehen als Modus der Welterschließung begreifen, um als mündige Bürger an einer Wissensgesellschaft wie der unseren partizipieren zu können (KMK 2005a; Mayer et al. 2004; Abd-El-Khalick et al. 2004; Harlen 1999; Harms et al. 2004; Mayer 2013; Schwartz et al. 2004). Außerdem kann die bloße Vermittlung des Ist-Standes von Wissen dem komplexen Prozess, der diesen Erkenntnissen vorausgeht, kaum dem Charakter von Wissenschaft gerecht werden (Crawford 2007; KMK 2005a; Zion et al. 2004). Deshalb sollen Lernende über fachwissenschaftliche Inhalte hinaus auch ein Verständnis für die Natur der Naturwissenschaften (*nature of science*) und die naturwissenschaftliche Erkenntnisgewinnung (*nature of inquiry*) entwickeln (Harms et al. 2004; iSTAR Assessment 2011; Mayer et al. 2004; Kremer und Mayer 2013; NRC 1996; Schwartz und Crawford 2006). Aus diesen Gründen sind Kompetenzen der Erkenntnisgewinnung integraler Bestandteil der naturwissenschaftlichen Grundbildung (*scientific literacy*) und damit ein international anerkanntes Ziel der Schulbildung (Bybee 2006). Daher sind sie auch in den Steuerungsdokumenten verschiedener Länder (z. B. Kanada: OME 2008a, 2008b; Großbritannien: DfES/QCA 2004; Australien: ACARA 2012 und USA: NGSS Lead States 2013; NRC 1996) und nicht zuletzt in den deutschen Bildungsstandards für den mittleren Schulabschluss (KMK 2005a, 2005b, 2005c) sowie den einheitlichen Prüfungsanforderungen der Abiturprüfung (KMK 2004) verankert und sollen auch in Oberstufenstandards entsprechendes Gewicht erhalten (Harms et al. 2004; Mayer et al. 2004, KMK 2005a, Kurth 2015; Stanat und Pant 2012).

Zusammenfassend steht der Biologie-Unterricht – speziell in der gymnasialen Oberstufe – vor der Aufgabe neben dem Aufbau eines grundlegenden Fachverständnisses auch Kompetenzen der naturwissenschaftlichen Erkenntnisgewinnung und ein angemessenes Bild von Wissenschaft und ihren Methoden¹ zu vermitteln. Hier wird das Forschende Lernen zur (gleichzeitigen) Förderung dieser Bereiche empfohlen (Mayer und Ziemek 2006). Allerdings wird die Wirksamkeit dieser Methode wegen der großen Schülerselbstständigkeit bzw. Offenheit sowie in Bezug auf

die vielfältigen Lernziele in Frage gestellt (Kirschner et al. 2006). Konsens scheint zu sein, dass das Forschende Lernen aufgrund seiner inhaltlichen Komplexität der Strukturierung oder Unterstützung durch die Lehrkraft bedarf, um einer Überforderung der Schüler vorzubeugen und somit erfolgreiches Lernen überhaupt erst möglich zu machen (Hmelo-Silver 2004). Bisher fehlen jedoch Studien, die untersuchen, wie das Forschende Lernen für die verschiedenen abhängigen Variablen lernwirksam unterstützt werden kann. Die vorliegende Arbeit sucht diese Lücke zu schließen. Erste Hinweise, dass etwaige Unterstützungsformate im Bereich des prozeduralen Wissens sowie des Methodenwissens benötigt werden, um das Forschende Lernen zu fördern, liegen bereits vor (Arnold et al. 2014). In der vorliegenden Studie werden nun zwei bestehende Unterstützungsformate zur Förderung des prozeduralen Wissens und des Methodenwissens für die Methode des Forschenden Lernens weiterentwickelt und für den Biologie-Unterricht der Oberstufe nutzbar gemacht und anschließend auf ihre Wirkung anhand einer quasi-experimentellen Interventionsstudie in der gymnasialen Oberstufe untersucht.

Wissenschaftliches Denken, Methodenwissen und Fachwissen

Das Wissenschaftliche Denken beschreibt die Fähigkeit, wissenschaftliche Untersuchungen verstehen, durchführen und kritisch reflektieren zu können, und es wird durch deklaratives Wissen (Methodenwissen und Fachwissen) beeinflusst (KMK 2004; Mayer 2007). Diese drei Konstrukte werden im Folgenden kurz vorgestellt.

Das Wissenschaftliche Denken wird als naturwissenschaftlicher Problemlöseprozess beschrieben (Abd-El-Khalick et al. 2004; Klahr und Dunbar 1988; Mayer 2007). Einer umfassenden Analyse von (Kompetenz-)Modellen und entsprechenden Testinstrumenten zufolge kann dieser Prozess in die Problemlöseprozeduren Fragen entwickeln, Hypothesen aufstellen, Untersuchungen planen und Daten auswerten unterteilt werden, die speziell für das Experimentieren wie folgt ausdifferenziert werden (Arnold 2015):

1. Fragestellungen entwickeln: abhängige und unabhängige Variable werden identifiziert und als Frage nach deren kausalem Zusammenhang formuliert (bspw. Chinn und Malhotra 2002; Germann et al. 1996; Germann und Aram 1996, 1996b; Harwood 2004, Hofstein et al. 2005, Lin und Lehman 1999),
2. Hypothesen aufstellen: abhängige und unabhängige Variable werden aus der Fragestellung aufgegriffen und als Vorhersage formuliert. Zudem wird die Hypothese begründet und es werden alternative Hypothesen formuliert

¹ Zu berücksichtigen ist, dass der Prozess der Erkenntnisgewinnung mit unterschiedlichen Untersuchungsmethoden durchlaufen werden kann (Wellnitz 2012; Wellnitz und Mayer 2013). Die vorliegende Arbeit beschäftigt sich exemplarisch mit dem Experiment. Es ist von wesentlicher Bedeutung in den Naturwissenschaften (Osborne et al. 2003) und dient der Untersuchung von kausalen Zusammenhängen, indem die unabhängige Variable variiert und die abhängige Variable beobachtet oder gemessen wird, während andere potentiell beeinflussende Variablen konstant gehalten bzw. kontrolliert werden (Mayer et al. 2008; Zion et al. 2004). Das Experimentieren stellt dabei spezifische Anforderungen an die Schüler (Mayer und Wellnitz 2014; Arnold et al. 2016).

- (bspw. Temiz et al. 2006; Tamir et al. 1992; Philips und German 2002; Klahr und Dunbar 1988),
3. Untersuchungen planen: die abhängige Variable wird operationalisiert und die unabhängige Variable variiert, sowie Störvariablen kontrolliert und darüber hinaus werden Messzeiten und Messwiederholungen festgelegt (bspw. Tobin und Capie 1982, Fraser 1980; Beaumont-Walters und Soyibo 2001; Klahr und Dunbar 1988) und schließlich
 4. Daten auswerten: Daten werden beschrieben und interpretiert, wobei die Interpretation in ihrer Sicherheit sowie die gesamte Methode einer kritischen Prüfung unterzogen wird. Ferner wird ein Ausblick gegeben (bspw. Tamir et al. 1992; Philips und German 2002; Chinn und Malhotra 2002; Klahr und Dunbar 1988, Meier und Mayer 2011).

Das Modell des Wissenschaftlichen Denkens beinhaltet nach Mayer (2007) zwei Arten deklarativen Wissens (Fach- und Methodenwissen) als Einflussfaktoren, die bereits bestätigt werden konnten (Arnold 2015). In unterschiedlichen Studien konnte gezeigt werden, dass die Schülerfähigkeiten im Bereich des Wissenschaftlichen Denkens hinter den erwarteten Leistungen zurückbleiben. Zum Beispiel haben Lernende Probleme beim Generieren von Forschungsfragen und können dabei nicht auf einen Faktor fokussieren (Hofstein et al. 2005; Kuhn und Dean 2005). Im Bereich der Hypothesen zeigen Lernende Schwierigkeiten, Hypothesen als Zusammenhang zwischen Variablen (de Jong und van Joolingen 1998) und generell als widerlegbare Vorhersage mit entsprechenden Alternativen zu formulieren (Klahr und Dunbar 1988; Klahr et al. 1993). Dies spiegelt sich dann auch in der Planung von Experimenten wieder, die häufig mit nur einer Variation der unabhängigen Variable und ohne Kontrollversuch auskommen (Hamman et al. 2008), keine Störvariablen berücksichtigen (Duggan et al. 1996) oder durch die Konfundierung der Variablen keine validen Aussagen zulässt (Chen und Klahr 1999). Zudem werden Experimente eher als Bestätigung der Hypothesen durchgeführt, also nur einmalig und ohne Wiederholung (Duggan und Gott 2000a; Lubben und Millar 1996). In der Datenauswertung finden sich Schwierigkeiten in dem Rückbezug zur Hypothese und der Begründung der Schlussfolgerung (Germann und Abram 1996b), Schlussfolgerungen kritisch reflektieren zu können und dabei bspw. Stichprobengröße, Repräsentativität und Validität zu berücksichtigen (Lubben und Millar 1996; Roberts und Gott 2004). In Studien, die speziell auf ältere Lerner fokussieren, konnten ähnliche Schwierigkeiten identifiziert werden (Tamir et al. 1992; Arnold et al. 2013). Zudem konnte in einer ergänzenden Videostudie (Arnold et al. 2014) festgestellt werden, dass die Lernenden häufig Probleme in der konkreten Umsetzung (bspw. bei der Festlegung von Messzeiten) haben, also nicht

wissen, *wie* vorzugehen ist und bspw. von Hilfestellungen auf prozeduraler Ebene profitieren könnten. Das Methodenwissen und das Fachwissen sollen als Einflussfaktoren des Wissenschaftlichen Denkens nun weiter ausgeführt werden.

Das Methodenwissen umfasst das Wissen über bzw. Verständnis für naturwissenschaftliche Methoden, deren Grenzen und Möglichkeiten sowie über Zweck und Funktion einzelner Aspekte wissenschaftlichen Arbeitens (Arnold 2015; Arnold et al. 2016). Davon ausgehend, dass wissenschaftliches Arbeiten mehr ist als die bloße Anwendung manueller Fähigkeiten und spezifisches Wissen benötigt (NRC 2012), beschreibt das Methodenwissen das Denken hinter dem Handeln (*thinking behind the doing*; Roberts 2001) und wird auch als das „Wissen, *warum*“ bezeichnet (bspw. *Warum braucht man eine Hypothese? Warum sollte die unabhängige Variable variiert werden?*; Glaesser et al. 2009a). Damit steht es in enger Verbindung zu den einzelnen Problemlöseprozeduren (siehe oben) und den Gütekriterien wissenschaftlichen Arbeitens (Objektivität, Reliabilität und Validität), die als Antwort auf die Frage „*Warum?*“ herangezogen werden können (Gott et al., o.J., Glaesser et al. 2009a). Ein Überblick über die in der Literatur aufgeführten Aspekte des Methodenwissens ist in Tab. 1 dargestellt.

Das hier beschriebene Methodenwissen ist angelehnt an das Konstrukt des *procedural understanding* (Gott und Duggan 1995) bzw. der *concepts of evidence* (Gott et al. o.J.; Gott und Roberts 2008), was von Glaesser und Kollegen (2009a) als deklaratives Wissen beschrieben wird, das dem prozeduralen Wissen, also dem „Wissen *wie*“ zugrunde liegt. Ferner argumentieren die Autoren, dass dieses Wissen eine notwendige Bedingung für das erfolgreiche Experimentieren darstellt und bestätigen dies in einer Studie (Glaesser et al. 2009b). In einer weiteren Studie konnte gezeigt werden, dass zwischen dem *procedural understanding*, das im *Evidence Test* (Roberts und Gott 2004) operationalisiert wurde, und der erfolgreichen Durchführung eines Experiments ein Zusammenhang besteht. Überschneidungen zwischen dem Methodenwissen und dem Konstrukt der *Nature of Science (NOS)* bestehen insofern, dass es darum geht, wie Wissenschaft funktioniert bzw. wie Wissenschaftler arbeiten. Glaesser und Kollegen (2009a) sehen das *procedural understanding* auch als Teil von *NOS* und diese Verbindung kann anhand einer Studie, die den Zusammenhang zwischen Vorstellungen zu *NOS* und dem Wissenschaftlichen Denken untersucht, unterstützt werden (Kremer et al. 2013). Stärker auf den Erkenntnisprozess zugeschnitten ist das Konstrukt der *Nature of Scientific Inquiry (NOSI)* (Neumann 2011; Schwartz et al. 2008; Lederman et al. 2014), das konkret die Vorstellungen zur naturwissenschaftlichen Erkenntnismethode beschreibt. Jedoch sind die Ansätze der *NOS* und der *NOSI* stark epistemologisch und soziologisch geprägt (Neumann und

Tab. 1 Aspekte des Methodenwissens

Aspekt	Quellen
Funktion des Experiments	Gott et al. o.J.; Osborne et al. 2003; Roberts, 2001; Wellnitz, 2012
Funktion von Hypothesen	Mayer und Ziemek 2006; Osborne et al. 2003
Funktion der unabhängigen Variable	Chen und Klahr 1999; Roberts 2001; Tamir et al. 1992
Funktion der abhängigen Variable	Chen und Klahr 1999; Roberts 2001; Tamir et al. 1992
Funktion der Planung	Lederman et al. 2014; Roberts 2001
Funktion der Operationalisierung von abhängigen Variablen	Gott et al. o.J.; Roberts und Gott 2003; Roberts 2001
Funktion der Variation von unabhängigen Variablen	Gott et al. o.J.; Roberts und Gott 2003; Roberts 2001
Funktion der Kontrolle von Störvariablen	Gott et al. o.J.; Roberts 2001
Funktion von Wiederholungen	Buffler et al. 2001; Gott et al. o.J.; Lubben und Millar 1996; Lubben et al. 2001; Osborne et al. 2003
Funktion der Trennung zwischen Beschreibung und Interpretation von Daten	Lederman et al. 2002; Lederman et al. 2014; Osborne et al. 2003
Funktion der Beschreibung von Daten	Lederman et al. 2002; Osborne et al. 2003
Funktion der Interpretation von Daten	Lederman et al. 2002; Lederman et al. 2014; Osborne et al. 2003
Funktion der kritischen Reflexion	Mayer und Ziemek 2006; Osborne et al. 2003
Gütekriterien Reliabilität, Validität, Objektivität	Gott et al. o.J.; Gott und Duggan 1995; Kirschner 1992; Osborne et al. 2003; Roberts 2001

Kremer 2013) und weniger spezifisch auf die prozedurale Ebene bezogen als das *procedural understanding*.

Im Bereich des Methodenwissens identifizierten Roberts und Gott (2004) anhand des *Evidence Test* Schwierigkeiten der Lernenden, bspw. in der Fähigkeit, die Validität und Aussagekraft von Experimenten mit Stichprobengröße und Repräsentativität zu begründen, warum Messwiederholungen sinnvoll sind. In einer Studie von Völzke, Arnold und Kremer (2013) wurden ebenfalls Defizite im Methodenwissen identifiziert, die sich auf die Planung von Experimenten auswirkten und Arnold et al. (2014) konnten zeigen, dass sich Lernende beim Experimentieren häufig nicht im Klaren darüber sind, *warum* sie beim Experimentieren manche Dinge tun. Die Autoren legen dar, dass eine Förderung des Methodenwissens nötig ist, um die Experimentierfähigkeit der Lernenden zu elaborieren. In Bezug auf den Erwerb von Methodenwissen wird argumentiert, dass solche Konzepte explizit gefördert werden sollten (Sandoval und Morrison 2003; Abd-El-Khalick und Lederman 2000; Akerston et al. 2000; Khishfe und Abd-El-Khalick 2002) und diese Förderung geeignet ist, das wissenschaftliche Denken zu unterstützen (Sandoval und Reiser 2004). Duschl (2000) beschreibt, dass die *NOS* dann explizit wird, wenn Lernende untersuchen, diskutieren, über Belege argumentieren und zwischen alternativen Erklärungen entscheiden. In einer Interventionsstudie von Chen und Klahr (1999) konnte gezeigt werden, dass sich Reflexionen bzw. Diskussionen über die Anwendung der einzelnen Schritte des Experiments und deren zugrundeliegende Logik (also die Erläuterung des *Warum*) positiv auf den Erwerb der Variablenkontrollstrategie auswirken und zudem für den Erwerb von Fachwissen förderlich sind.

Das Fachwissen schließlich ist das Wissen über die jeweiligen biologischen Phänomene, die den Gegenstand der Untersuchung bilden. Das Fachwissen spielt beim Wissenschaftlichen Denken bzw. dem Forschenden Lernen zwei verschiedene Rollen. Einerseits beschreibt es das Wissen über bzw. Verständnis für Fakten, Konzepte, Gesetze und Theorien, das für die Lösung des Problems notwendig ist und beeinflusst so die Fähigkeit des Wissenschaftlichen Denkens (Mayer 2007). Dieser Zusammenhang konnte bereits in mehreren Studien gezeigt werden (bspw. Duggan et al. 1996; Solano-Flores et al. 1999; Glaesser et al. 2009a; Wellnitz und Mayer 2013; Hof 2011; Hammann et al. 2007). Andererseits wird das wissenschaftliche Arbeiten häufig auch genutzt, um Fachwissen zu erarbeiten und zu erlernen. Diesbezüglich ist die Studienlage uneindeutig. Während in der Meta-Analyse von Dochy und Kollegen (2003) das wissenschaftliche Arbeiten tendenziell negative Effekte auf den Erwerb von Fachwissen zeigt, stellen die Autoren fest, dass der Effekt stark durch das Expertise-Level der Lernenden moderiert wird und zudem, dass das Behalten durch das wissenschaftliche Arbeiten positiv beeinflusst wird. In der Meta-Analyse von Furtak und Kollegen (2009) hingegen zeigt sich ein positiver Effekt auf den Erwerb von Fachwissen. In dieser Rolle, also als abhängige Variable des wissenschaftlichen Arbeitens im Rahmen des Forschenden Lernens, wird das Fachwissen in der vorliegenden Studie verstanden. Im Folgenden wird näher auf das Forschende Lernen eingegangen.

Forschendes Lernen

Das Forschende Lernen wird als Methode zur Förderung des Wissenschaftlichen Denkens empfohlen (Mayer und Ziemek 2006; KMK 2004). In Anlehnung an Hmelo-Silver und Kollegen (2007), Mayer und Ziemek (2006) sowie Gijbels und Kollegen (2005) wird in dieser Arbeit davon ausgegangen, dass sich das Forschende Lernen durch zehn Merkmale auszeichnet: Das Forschende Lernen ist 1) eine konstruktivistische und kollaborative und 2) problemorientierte und schülerzentrierte Unterrichtsmethode, die 3) dem Erwerb von Fachwissen, wissenschaftsmethodischen Kompetenzen und Problemlösefähigkeiten dienen kann. Beim Forschenden Lernen werden 4) (von den Lernenden) aus authentischen Problemen Fragestellungen und Hypothesen abgeleitet, Untersuchungen (z. B. Experimente) geplant und durchgeführt und Daten ausgewertet, 5) die entsprechenden Versuche (Experimente) sind ergebnisoffen angelegt und 6) die Lernenden entscheiden selbst über die Ausgestaltung des Versuchs (Experiments). Zudem werden 7) Versuche (Experimente) mehrmals durchgeführt und ausgewertet, wobei das Vorgehen anschließend kritisch reflektiert und die Sicherheit der gewonnenen Ergebnisse diskutiert und ein Ausblick auf weitere Versuche gegeben wird. Beim Forschenden Lernen 8) arbeiten die Lernenden weitestgehend selbstständig in kleinen Gruppen, und 9) generieren und konstruieren Wissen und Fähigkeiten selbst, wobei 10) die Lehrperson in den Hintergrund tritt und als Moderator und Hilfegeber die Entdeckungsprozesse der Lernenden begleitet (Arnold 2015). Die Lernförderlichkeit dieser Methode konnte in den Bereichen Wissenschaftliches Denken (Dochy et al. 2003; Furtak et al. 2009; Blanchard et al. 2010; Hof 2011), Methodenwissen bzw. *NOS*² (Furtak et al. 2009; Blanchard et al. 2010; Glaesser et al. 2009a) und Fachwissen (Dochy et al. 2003; Hattie 2009; Furtak et al. 2009; Blanchard et al. 2010; Hof 2011) bereits (in Teilen) belegt werden. Allerdings ist die Studienlage uneindeutig (bspw. Dochy et al. 2003), was nicht zuletzt darauf zurückzuführen ist, dass „Forschendes Lernen“ unterschiedlich gefasst wird und die Studien sich in den abhängigen Variablen und deren Operationalisierung teils stark unterscheiden (Hattie 2009): Studien zeigen, dass die Effektivität des Forschenden Lernens bspw. von der adäquaten Umsetzung (Blanchard et al. 2010) bzw. dem Grad der Offenheit (Hof 2011; Kirschner et al. 2006; Klahr und Nigam 2004; Dean und Kuhn 2007) abhängt. Das Forschende Lernen kann in vielerlei Hinsicht als „offen“ beschrieben werden,

bspw. in Bezug auf Fachinhalt, Phasen, Strategie, Methode, Lösung oder Lösungswege (Priemer 2011). Es wird jedoch argumentiert, dass die Offenheit, vor allem bei schwachen Schülern, zu hoher kognitiver Belastung führt, und daher nicht effektiv sein kann (Kirschner et al. 2006). Andererseits kann die direkte Instruktion die komplexe Natur der Naturwissenschaften nicht authentisch abbilden (Chinn und Malhotra 2002) und ist nicht konsistent mit konstruktivistischen Sichtweisen von Lernen (Hmelo-Silver et al. 2007). Daraus wird ersichtlich, dass die Unterstützung der Lehrkraft beim Forschenden Lernen ein zentraler Faktor ist. Der Ansatz des angeleiteten Forschenden Lernens (*Guided inquiry*; Furtak 2006) scheint hier eine Lösung darzustellen, denn er kombiniert die Essenz des offenen Forschenden Lernens mit instruktionaler Unterstützung. Durch die zeitweise Unterstützung oder ggf. Anleitung wird dann entsprechend der Grad der Offenheit reduziert. Allerdings stellt es eine große Herausforderung für die Lehrkräfte dar, darüber zu entscheiden, wann welche Unterstützungen notwendig sind und wie trotzdem der forschende Charakter aufrechterhalten werden kann (Crawford 2000; Furtak 2006). Daher ist es sinnvoll, zumindest teilweise standardisierte, Lernunterstützungen für das Forschende Lernen zu entwickeln, empirisch zu testen und so für die Förderung bereitzustellen.

Lernunterstützungen beim Forschenden Lernen

Komplexe Aufgaben, wie es naturwissenschaftliche Problemlöseaufgaben im Rahmen des Forschenden Lernens sind, können Schüler schnell überfordern (Schmidt-Weigand et al. 2008). Dies lässt sich mit der *Cognitive Load Theory* (Sweller et al. 1998), die davon ausgeht, dass das Arbeitsgedächtnis begrenzt ist, erklären. Demzufolge sollte es das Ziel sein, Materialien und die geforderten Aktivitäten so zu gestalten, dass die durch die Arbeitsaufträge bzw. Materialien verursachte Belastung (*extraneous Cognitive Load*) reduziert wird, und so mehr Kapazität für die Konstruktion von Schemata und damit zum Lernen (*germane Cognitive Load*; Schmidt-Weigand et al. 2008; Sweller et al. 1998) aufgewendet werden kann.

Ein wichtiger Ansatz im Bereich der Lernunterstützungen ist der der *scaffolds* (zu Deutsch: Gerüste). Der Begriff wird teils sehr breit und synonym mit jeglicher Art von Unterstützung gebraucht. Dabei geht jedoch der wahre Charakter von *scaffolds* verloren, denn eine andauernde Diagnose, die Anpassung der Unterstützung, sowie das Ausschleichen der Unterstützung (analog zum Abbau des Gerüsts) sind wichtige Faktoren des *scaffolding* im eigentlichen Sinne (Puntambekar und Hübscher 2005). Saye und Brush (2002) unterscheiden zwei Arten von Unterstützungen: *soft scaffolds* und *hard scaffolds*. Als *soft scaffolds*

² Methodenwissen kann als Teil der *Nature of Science* gesehen werden (Gott und Roberts 2008; Arnold et al. 2014). Da über das Methodenwissen selbst nur wenige Befunde vorliegen, werden folgend Befunde zu Methodenwissen im Speziellen und zur *Nature of Science* im Allgemeinen gemeinsam dargestellt.

beschreiben sie dynamische und situative Unterstützungsmaßnahmen der Lehrkräfte (*scaffolds* i. e. S.). Dies verlangt auf Seiten der Lehrkraft viel Aufmerksamkeit und situative Diagnostik von Verständnisproblemen und entsprechend den abgestimmten Einsatz von Hilfestellungen. *Hard scaffolds* hingegen sind statische Unterstützungen, die typische Schülerprobleme antizipieren und vorbereitet werden können (Saye und Brush 2002). Sie werden häufig in computerbasierten Experimentier-Umgebungen genutzt (Wichmann und Leutner 2009; Künsting et al. 2010). In der vorliegenden Studie liegt der Schwerpunkt auf den *hard scaffolds* und deren Verwendung in Hands-on-Experimenten. Dies liegt darin begründet, dass sie einerseits fertige Materialien darstellen, die die Lehrenden entlasten (Hmelo-Silver 2006) und es ihnen zudem ermöglichen, sich weiter aus der schüler selbstständigen Arbeit herauszuziehen (siehe Kapitel „Forschendes Lernen“), und andererseits eignen sie sich aus methodischen Gründen besser als *soft scaffolds*, da sie standardisiert und so systematisch in ihrer Wirkung untersucht und dann weiter verwendet werden können. Dabei ist der Anspruch, möglichst viele der positiven Aspekte von *soft scaffolds* in die Lernunterstützungen zu integrieren. Folgend werden zwei Formate als Lernunterstützungen vorgestellt, die sich im Rahmen des Forschenden Lernens eignen: gestufte Lernhilfen und diskursiv-reflexive Szenarien.

Gestufte Lernhilfen

Gestufte Lernhilfen sind eine Weiterentwicklung von Lösungsbeispielen (*worked-out examples*) (Hänze et al. 2010; Forschergruppe Kassel 2007). Zweitere haben im Bereich des Problemlösens eine lange Tradition zur Förderung von kognitiven Fähigkeiten (Atkinson et al. 2000). Sie kommunizieren den Prozess und geben konkrete Hilfestellungen zur Lösung der Aufgabe. Lösungsbeispiele bestehen aus der Problemformulierung, Lösungsschritten und der finalen Lösung (Renkl 2005). Sie können zwar ausgeschlossen werden, jedoch findet keine Diagnose und entsprechende Anpassung der Unterstützung statt. Die Wirksamkeit von Lösungsbeispielen gegenüber dem offenen Problemlösen konnte bspw. von Atkinson und Kollegen (2000) nachgewiesen werden.

Bei den gestuften Lernhilfen, die ähnlich wie Lösungsbeispiele aufgebaut sind, werden Lösungshinweise jedoch schrittweise gegeben und die Schüler können nach Wunsch oder Bedarf darauf zurückgreifen (Schmidt-Weigand et al. 2008; Schmidt-Weigand et al. 2009). Somit werden Diagnose und Ausschleichen in Schülerhand gegeben. Im Vergleich zu konventionellen Lösungsbeispielen konnte im Bereich physikalischen Fachwissens bereits gezeigt werden, dass gestufte Lernhilfen lernförderlicher sein können (Schmidt-Weigand et al. 2008). Im Bereich

naturwissenschaftlicher Untersuchungen (Physik) zeigte sich, dass mit Lösungsbeispielen und gestuften Lernhilfen mehr Lösungsschritte erinnert und höhere Leistungen im anschließenden Fachwissenstest erzielt werden als beim offenen Problemlösen (Schmidt-Borcherding et al. 2013). Bislang fehlen Instruktionsstudien, die die Wirksamkeit von gestuften Lernhilfen beim wissenschaftlichen Arbeiten auf den Erwerb von Wissenschaftlichem Denken oder Methodenwissen differenziert untersuchen.

Diskursiv-reflexive Szenarien (*Concept Cartoons*)

Als Format zur Anregung von Diskussionen über das Methodenwissen (siehe Kapitel „Wissenschaftliches Denken, Methodenwissen und Fachwissen“) wurden in dieser Arbeit *Concept Cartoons* (Naylor und Koegh 1999) gewählt. Es handelt sich um *hard scaffolds*, die der Sprech-anregung dienen. *Concept Cartoons* beinhalten die Darstellung verschiedener Figuren (Cartoons), die über Sachverhalte diskutieren und dienen als Impulse zur Äußerung von Vorstellungen, deren Diskussion und somit der Ko-Konstruktion von Wissen (Keogh 1999). Özmen und Kollegen (2012) überprüften die Wirksamkeit von *Concept Cartoons* auf den Erwerb von Fachwissen innerhalb von Laboraktivitäten und zeigten damit signifikant höhere Lernzuwächse gegenüber einem Vergleich ohne diese Lernunterstützung. Über die Effektivität von *Concept Cartoons* zur Förderung von Wissenschaftlichem Denken und Methodenwissen liegt nach derzeitigem Stand nur eine Studie vor, die die Wirksamkeit der *Concept Cartoons* in Bezug auf das Methodenwissen nahelegt und zeigt, dass sie sich partiell auf die Performanz beim Experimentieren auswirken können (Kuhn et al. 2000).

Ziel der Studie und Hypothesen

Bisher wurde die Wirksamkeit beider Lernunterstützungen (gestufte Lernhilfen und *Concept Cartoons*) sowie deren Kombinationen nicht systematisch im Zusammenhang mit dem Forschenden Lernen in Bezug auf den Erwerb von Wissenschaftlichem Denken, Methodenwissen und Fachwissen sowie die empfundene kognitive Belastung untersucht. Ziel der vorliegenden Studie ist es daher zu prüfen, inwiefern deren Nutzung und somit die gezielte Förderung des prozeduralen Wissens bzw. des Methodenwissens sich eignet, um den Erwerb von Wissenschaftlichem Denken, Methodenwissen und Fachwissen zu fördern und die kognitive Belastung zu reduzieren. Im Folgenden wird immer von der Wirkung der einzelnen Unterstützungsformate (gestufte Lernhilfen und diskursiv-reflexive) sowie deren Kombination ausgegangen. Die Kombination wird untersucht, weil geprüft werden soll, ob die beiden Unterstützungen

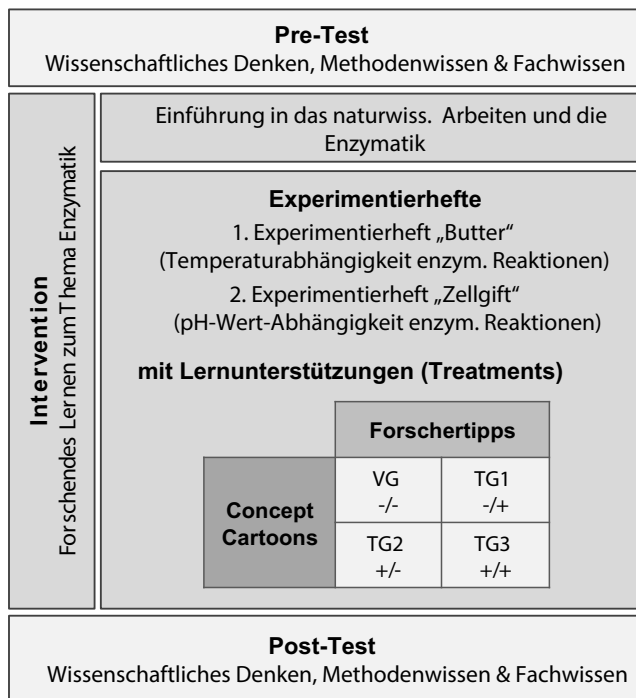


Abb. 1 Übersicht Studiendesign

gen, die jeweils unterschiedliche Wissensarten (prozedurales Wissen und deklaratives Methodenwissen) ansprechen, sich sinnvoll ergänzen können. Dies liegt nahe, weil davon ausgegangen werden kann, dass die beiden Wissensarten zusammenhängen und sich somit positiv beeinflussen können (siehe Kapitel „Wissenschaftliches Denken, Methodenwissen und Fachwissen“). Darüber hinaus kann aus *Cognitive Load*-Sicht argumentiert werden, dass eine Entlastung durch Unterstützung im einen Bereich Arbeitsgedächtniskapazitäten für den anderen Bereich zur Verfügung stellt und somit die Kombination aus beiden den Effekt erhöhen könnte. Allerdings ist auch denkbar, dass ein Mehr an Materialien oder Informationen in Form von Lernunterstützungen zusätzliche Kapazitäten bindet.

Interventionsstudien, die die Wirksamkeit von gestuften Lernhilfen beim wissenschaftlichen Arbeiten auf den Erwerb von Wissenschaftlichem Denken beim Forschenden Lernen untersuchen, sind derzeit nicht bekannt. Es wird jedoch vermutet, da dieses Instruktionsformat allgemein das prozedurale Wissen fördern kann (Atkinson et al. 2000), dass es auch in diesem Bereich lernförderlich ist. Darüber hinaus kann aufgrund der oben angeführten Befunde erwartet werden, dass sich die Lernunterstützungen sowie deren Kombination positiv auf den Erwerb von Wissenschaftlichem Denken auswirken ($H_{\text{Wissenschaftliches Denken}}$).

Für die Lernförderlichkeit von gestuften Lernhilfen in Bezug auf das Methodenwissen liegen ebenfalls keine Ergebnisse vor. Es wird jedoch erwartet, dass sich die Erläuterung der einzelnen Schritte des Prozesses wissenschaftli-

chen Arbeitens positiv auf das Methodenwissen auswirken. Hier wird zusätzlich auf Grundlage der oben genannten Studien erwartet, dass sich die Lernunterstützungen und die Kombination aus diesen positiv auf den Erwerb von Methodenwissen auswirken ($H_{\text{Methodenwissen}}$).

In Bezug auf Fachwissen konnte die Lernförderlichkeit von gestuften Lernhilfen (im Kontext wissenschaftlichen Arbeitens) mit Bezug zum Fachwissen bereits tendenziell, wenn auch nicht statistisch signifikant, belegt werden (Schmidt-Borcherding et al. 2013). Auch für fachwissensbezogene *Concept Cartoons* konnte im Rahmen von Laboraktivitäten gezeigt werden, dass sie sich in Bezug auf das Fachwissen positiv auswirken (Özmen et al. 2012). Der Einfluss dieser Lernunterstützungen mit methodenbezogenen Inhalten ist noch nicht untersucht. Hier kann aber erwartet werden, dass durch die Lernunterstützungen im Methodenbereich kognitive Kapazitäten für das Erlernen von Fachwissen frei werden (Hof 2011). Daher ist zu erwarten, dass sich die Lernunterstützungen positiv auf den Erwerb von Fachwissen auswirken ($H_{\text{Fachwissen}}$).

Weiterhin soll untersucht werden, ob die Lernunterstützungen geeignet sind, den *Cognitive Load* zu reduzieren. Es wird davon ausgegangen, dass durch die gestuften Lernhilfen die Aufgaben stärker strukturiert und die geförderten Aktivitäten vereinfacht werden können und somit der *extraneous Cognitive Load* reduziert wird. Durch die Bearbeitung der *Concept Cartoons* wird für den Problemlöseprozess relevantes Hintergrundwissen aktiviert, das die Bearbeitung der Problemlöseaufgaben vereinfachen und so ebenfalls den *extraneous Cognitive Load* reduzieren kann. Daher wird erwartet, dass die Schüler mit Lernunterstützungen weniger kognitive Belastung berichten ($H_{\text{Cognitive Load}}$).

Design & Methodik

Zur Prüfung der Hypothesen wurde eine quasi-experimentelle Interventionsstudie im 2×2 -faktoriellen Design mit Pre- und Post-Test in neun Einführungskursen der gymnasialen Oberstufe (11. Jgst.) zum Thema Enzymatik durchgeführt ($N = 220$). Die Interventionsstudie dauerte ca. 8 Wochen und wurde als Feldstudie im Regelunterricht von den jeweiligen Lehrkräften durchgeführt (Abb. 1).

Während der Studie arbeiteten die Lernenden gemäß der obigen Beschreibung des Forschenden Lernens selbstständig in Gruppen von zwei bis drei Personen („Forscherteams“). Die Intervention begann mit einer Einführung in das naturwissenschaftliche Arbeiten und die Enzymatik, die insgesamt ca. drei Unterrichtsstunden in Anspruch nahmen und anhand von Arbeitsblättern³ angeleitet wurden. Während der Einführung wurde das Experiment inkl. abhängige

³ Die kompletten Materialien sind in Arnold 2015 dokumentiert.

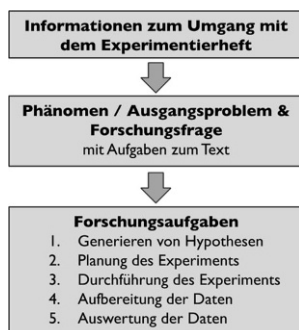


Abb. 2 Struktur der Experimentierhefte

und unabhängige Variable sowie die Gütekriterien wissenschaftlichen Arbeitens (Objektivität, Reliabilität und Validität) beschrieben. Außerdem wurden die Teilprozesse (Fragestellung, Hypothese, Planung, Durchführung und Datenauswertung) mit den entsprechenden Aspekten behandelt (siehe Kapitel „Wissenschaftliches Denken, Methodenwissen und Fachwissen“). Bei der Einführung in die Enzymatik wurde die allgemeine Funktion von Enzymen beschrieben und es wurden Begriffe wie Biokatalysator, Substrat, Produkt und Aktivierungsenergie geklärt. Dadurch erhielten die Lernenden das für die selbstständige Bearbeitung der folgenden Materialien notwendige Wissen. Anschließend folgte die schüler selbstständige Bearbeitung von zwei Experimentierheften zum Thema Enzymatik mit unterschiedlichen Lernunterstützungen (unabhängige Variablen). Hier bearbeiteten die Gruppen jeweils ein Experimentierheft gemeinsam und wechselten selbstständig nach Abschluss des ersten Experimentierhefts zum zweiten. Die beiden Experimentierhefte waren strukturell gleich aufgebaut (Abb. 2), hatten jedoch unterschiedliche thematische Schwerpunkte.

Bei der Unterrichtseinheit wurde darauf geachtet, dass die Lernenden selbstständig mit den Materialien arbeiten und die Lehrkräfte wirklich nur als Moderatoren und Hilfengeber tätig sein konnten (Theyßen 2014). Die Lehrkräfte wurden entsprechend in mehreren Besprechungen in die Studie und ihre Rolle eingeführt. Der Unterricht wurde in allen Gruppen regelmäßig hospitiert, und die Einhaltung der vereinbarten Richtlinien kann bestätigt werden. Regelmäßige Gespräche und Pausenkonferenzen dienten dazu, den zeitlichen Verlauf zu koordinieren. Für Gruppen, die entsprechend schneller mit der Bearbeitung fertig waren, waren zusätzliche Aufgaben intendiert, jedoch ergaben sich keine zeitlichen Differenzen im Abschluss der beiden Experimentierhefte. Dies kann ggf. darauf zurückgeführt wer-

den, dass die Lernenden mit den Lernunterstützungen zwar zusätzlichen Leseaufwand hatten, aber dann aufgrund der Unterstützung zügiger arbeiten konnten.

Pre- und Post-Test wurden flankierend direkt vor bzw. nach der Unterrichtseinheit durchgeführt.

Unabhängige Variablen: Lernunterstützungen

Die beiden Unterstützungsformate diskursiv-reflexive Szenarien (*Concept Cartoons*) und die prozessunterstützenden gestuften Lernhilfen (*Forschertipps*) wurden im 2×2 -Design, wie in Abb. 1 dargestellt, auf drei Treatmentgruppen und eine Vergleichsgruppe verteilt (Tab. 2). Der Unterricht wurde in allen Gruppen gleich gestaltet und es lagen die gleichen Experimentierhefte zugrunde, die sich lediglich in den Lernunterstützungen unterschieden. Die Vergleichsgruppe erhielt keine zusätzliche Unterstützung zur Bearbeitung der Experimentierhefte. Die Zeit zur Bearbeitung wurde in allen Gruppen gleich gehalten.

Die prozessunterstützenden gestuften Lernhilfen (*Forschertipps*) fanden während der Forschungsaufgaben Anwendung. Sie beinhalteten Erläuterungen der jeweiligen Teilprozesse sowie Lösungsbeispiele für die Aufgaben, mit dem Ziel die einzelnen Schritte des Erkenntnisprozesses zu verdeutlichen und die Schüler in der Anwendung zu unterstützen (prozedurales Wissen). Bei den *Forschertipps* handelt es sich um gestufte Lernhilfen, die aus zwei Seiten bestehen: auf der ersten Seite findet sich ein Denkanstoß mit Lösungshinweisen und auf der zweiten Seite eine Beispiellösung mit Erklärung (Abb. 3).

Auf die *Forschertipps* wurde bei den jeweiligen Forschungsaufgaben hingewiesen (Verlinkungen), damit die Schüler wussten, wann sie auf welchen Tipp zugreifen konnten. Die Schüler der entsprechenden Treatmentgruppen erhielten pro Gruppe einen Umschlag je Experimentierheft mit je 15 kleineren Umschlägen, die die einzelnen *Forschertipps* enthielten. Diese Umschläge waren beschriftet und nummeriert, sodass sie den einzelnen Teilkompetenzen und Kompetenzaspekten zugeordnet werden konnten. Dadurch, dass die Umschläge verklebt waren, war es möglich, deren Nutzung nachzuvollziehen. Die Auszählung der *Forschertipps* ergab, dass sie in den entsprechenden Gruppen eher selten genutzt wurden. Während in Treatmentgruppe 1, die nur die *Forschertipps* zur Verfügung hatte, im ersten Experimentierheft ca. 20 % und im zweiten Experimentierheft ca. 23 % der Hilfen genutzt wurden, hat

Tab. 2 Übersicht über die Lernunterstützungen

	Concept Cartoons	Forschertipps
Format	Diskursiv-reflexive Szenarien	Gestufte Lernhilfen mit Lösungsbeispielen
Lerninhalt	Methodenwissen	Wissenschaftliches Denken
Unterstützte Wissensart	Deklaratives Wissen	Prozedurales Wissen

Hypothese	Forschertipp 1: unabhängige Variable
<p>Was genau wollt ihr untersuchen?</p> <p>Schaut euch die Forschungsfrage erneut an und überlegt, den Einfluss welches Faktors (unabhängige Variable) ihr in eurem Experiment untersuchen wollt. Welcher Faktor könnte eine Ursache für Veränderungen oder Unterschiede sein?</p>	
Hypothese	Forschertipp 1: unabhängige Variable
<p>Beispiellösung</p> <p>Ein möglicher Faktor, dessen Einfluss ihr untersuchen könnt, ist die Temperatur. Die Temperatur ist eine mögliche unabhängige Variable.</p>	

Abb. 3 Forschertipp „Hypothese – unabhängige Variable“

die Treatmentgruppe 3, der beide Lernunterstützungen zur Verfügung standen, ca. 12 % bzw. 18 % der Hilfen genutzt (Arnold 2015).

Die diskursiv-reflexiven Szenarien (*Concept Cartoons*; Abb. 4) wurden in Anlehnung an die aus der Theorie hergeleiteten Aspekte des Methodenwissens (Tab. 1) konstruiert. Sie dienen dazu, Diskussionen über Sinn und Zweck einzelner Schritte des wissenschaftlichen Erkenntnisprozesses zu initiieren und somit die Aufmerksamkeit auf die hinter dem Prozess liegenden Konzepte (Methodenwissen) zu lenken.

Die Bearbeitung der Cartoons wurde ebenfalls ausgewertet. Es konnte gezeigt werden, dass ca. 64 % der von den Lernenden notierten Begründungen für die jeweiligen Schritte des Experiments einen (indirekten) Bezug zu den Gütekriterien wissenschaftlichen Arbeitens aufwiesen und somit den intendierten Begründungen entsprachen (Arnold et al. 2016).

Abhängige Variablen und Instrumente

Die abhängigen Variablen Wissenschaftliches Denken, Methodenwissen und Fachwissen (Enzymatik) wurden in schriftlichen Leistungstests operationalisiert und entspre-

chend vorgetestet (Arnold 2015). Die finalen Tests wurden vor und nach der Intervention als Pre- bzw. Post-Test eingesetzt. Die Test-Kennwerte sind in Tab. 3 dargestellt. Dabei ist zu berücksichtigen, dass Winsteps (Linacre 2011) zwei Werte für Reliabilitäten ausgibt: für Items und Personen. Der erste Wert, die Item-Reliabilität, ist ein Index für die Reproduzierbarkeit der Item-Reihenfolge und hat kein Äquivalent in der klassischen Testtheorie (Bond und Fox 2012). Der Reliabilitäts-Wert für Personen hingegen ist die sog. Test-Reliabilität (Bond und Fox 2012; Linacre 2011). Er ist analog zu Koeffizienten der internen Konsistenz, wie bspw. Cronbachs α , der ebenfalls angegeben ist (Linacre 2011). Für diesen Wert gibt Winsteps (Linacre 2011) untere („real“) and obere („model“) Grenzen der echten Reliabilität an (Linacre 2011). Die Rasch-Reliabilitäten sind die konservativeren Werte. Zum Vergleich von Gruppen ist eine Reliabilität von 0,5 ausreichend (Lienert und Raatz 1998).

Der Test zum Wissenschaftlichen Denken wurde im offenen Antwortformat verfasst und bestand aus insgesamt sechs Items (Arnold 2015). Er enthielt je zwei Items zu den Teilkompetenzen „Hypothese“, „Planung“ und „Daten“. Da es sich um offene Aufgaben handelt, wurden die Schülerantworten mittels Codieranleitung bewertet (Cohens κ : 0,89–0,94 je nach Teilkompetenz).

Der Test zum Methodenwissen wurde in Anlehnung an die Aspekte des Methodenwissens (Tab. 1) entwickelt. Die finale Version des Tests bestand aus 15 Multiple Choice-Items. Ein Beispiel-Item ist in Abb. 5 dargestellt.

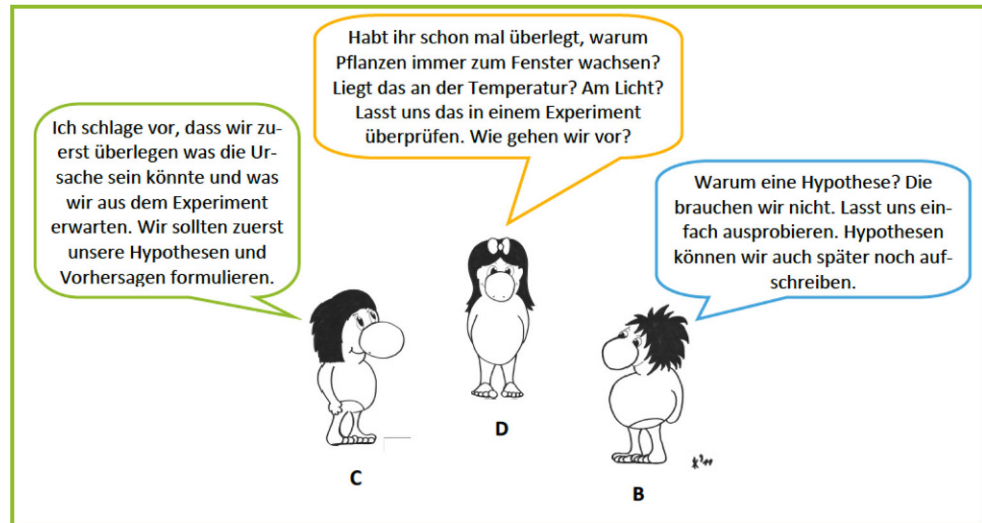
Das Fachwissen zum Thema Enzymatik wurde in Anlehnung an die Unterrichtsmaterialien der Intervention operationalisiert (Arnold 2015). Hier wurde das Fachwissen über die Katalase-Reaktion, die Lipolyse, sowie die Temperatur- bzw. pH-Wert-Abhängigkeit enzymatischer Reaktionen erfasst. Ein Beispiel-Item ist in Abb. 6 dargestellt.

Außerdem wurde die kognitive Belastung (*Cognitive Load*) jeweils nach Beendigung der beiden Experimentierhefte erfasst. Hierbei wurde auf die Skalen von Küsting (2007) in gekürzter und leicht veränderter Form zurückgegriffen.

Stichprobe

Insgesamt haben 220 Schüler an der Intervention teilgenommen. Die Zahl der Schüler, die sowohl am Pre-Test als auch am Post-Test teilgenommen haben betrug je Gruppe zwischen 40 und 56 Personen. Zu Beginn der Intervention waren die Schüler zwischen 15 und 19 ($M = 16,92$) Jahre alt; 60 % der Probanden waren weiblich. Die Lernenden hatten zu Beginn der Studie noch keinen Unterricht im Bereich der Enzymatik. Zur Kontrolle der Ausgangsbedingungen wurden die Pre-Test-Daten der jeweiligen Treatmentgruppen mittels Kruskal-Wallis-Test vergli-

Abb. 4 Concept Cartoon „Wozu braucht man eine Hypothese?“ (Zeichnung: Kathrin Ziepprecht)



Cartoon 1: Wozu braucht man eine Hypothese?

Beantwortet die Frage, indem ihr die Aussagen diskutiert. Bezieht dazu die Gütekriterien wissenschaftlichen Arbeitens ein. Begründet eure Antwort. Geht dazu so vor:

1. Jeder in der Gruppe sagt seine Meinung.
2. Ihr diskutiert über die unterschiedlichen Meinungen.
3. Ihr schreibt euren Konsens auf.

Tab. 3 Testkennwerte Tests Wissenschaftliches Denken, Methodenwissen und Fachwissen

Kriterium	Einheit	Soll	Test		
			Wissenschaftliches Denken	Methodenwissen	Fachwissen
Reliabilität	Item Reliability	Nahe 1,0 (max; nach Linacre 2011)	0,99	0,99	0,98
	Cronbach α	>0,5 (nach Lienert und Raatz 1998)	0,75	0,54	0,61
	Person Reliability (real-model)		0,63–0,68	0,53–0,57	0,48–0,51
Validität	MNSQ (In-Fit)	Nahe 1,0; $0,5 < \text{MNSQ} < 1,5$ (nach Wright und Linacre 1994)	0,87–1,09 ($M = 1,00$)	0,82–1,1 ($M = 1,00$)	0,90–1,16 ($M = 1,00$)
	MNSQ (Out-Fit)		0,62–1,66 ($M = 0,99$)	0,78–1,51 ($M = 1,32$)	0,86–1,44 ($M = 1,02$)
Objektivität	Cohens κ	>0,81 (nach Landis und Koch 1977)	0,89–0,94	/	/

chen und es konnten keine signifikanten Unterschiede zwischen den Gruppen festgestellt werden (*Wissenschaftliches Denken*: $H(3) = 3,40$, $p = 0,334$; *Methodenwissen*: $H(3) = 1,27$, $p = 0,735$; *Fachwissen*: $H(3) = 1,57$, $p = 0,665$). Da dennoch tendenzielle Unterschiede zwischen den Gruppen bestehen, werden die Residualgewinne berechnet und als Grundlage der folgenden Intergruppenvergleiche verwendet, um diese Unterschiede zu minimieren (Walpuski und Sumfleth 2007).

Datenanalyse

Die Datenanalyse erfolgte mittels *Item-Response-Theory (IRT)* mit dem Programm Winsteps (Linacre 2011) und Varianzanalysen (ANOVA) mit dem Programm SPSS. Auf die IRT wurde zur Testanalyse und Datenaufbereitung aus ver-

schiedenen Gründen zurückgegriffen. Erstens aufgrund der für das Rasch-Modell spezifischen Objektivität, die es erlaubt, die Limitationen, die durch die klassische Testtheorie entstehen, zu überwinden (Fox und Jones 1998). Dies ist teils dadurch gewährleistet, dass das Rasch-Modell Itemschwierigkeiten und Personenfähigkeiten nicht konfundiert (Fox und Jones 1998). Die Nutzung des Rasch-Modells, wenn die Daten zum Modell passen, gewährleistet somit, dass *Person-Measures* nicht von dem Itemsatz abhängen, den die Person beantwortet hat bzw. dass die Einschätzung der Itemschwierigkeit nicht von der Personengruppe abhängt, die sie beantwortet hat. Zweitens stellt die Rasch-Skalierung intervallskalierte Daten bereit, die es erlauben, parametrische Tests durchzuführen (Boone et al. 2014). Rohdaten von Tests oder Fragebögen hingegen können nonlinear sein und daher die Voraussetzung von parametrischen

Abb. 5 Beispiel-Item Methodenwissen (Aspekt „Zweck von Hypothesen“; Arnold 2015)

Welchen Zweck erfüllen Hypothesen in einem naturwissenschaftlichen Experiment **nicht**? Kreuze an.

Hypothesen lenken die Planung und Auswertung von Experimenten.

Hypothesen legen den Ausgang des Experiments fest.

Hypothesen dienen dem systematischen Ausschluss unwahrscheinlicher Erklärungen.

Hypothesen dienen der systematischen Findung von Erklärungsmöglichkeiten.

Abb. 6 Beispiel-Item Fachwissen (Arnold 2015)

Was passiert bei der **Lipolyse**? Kreuze an.

Lipolyse ist der Vorgang, bei dem...

... Fette durch Lysozyme zu Glycerin und Fettsäuren unter Bindung von Wasser gespalten werden.

... Fette durch Lipasen zu Glucose und Fettsäuren unter Bindung von Wasser gespalten werden.

... Fette durch Lipasen zu Glycerin und Fettsäuren unter Abspaltung von Wasser gespalten werden.

... Fette durch Lipasen zu Glycerin und Fettsäuren unter Bindung von Wasser gespalten werden.

Tests verletzen (Boone et al. 2014). Drittens eignet sich das Rasch-Modell bei Datensätzen mit fehlenden Daten. D. h., dass Schüler, wenn sie ein Item nicht bearbeitet haben, dennoch sinnvoll in ihrer Fähigkeit eingeschätzt werden können (Boone und Scantlebury 2006). Der vierte Grund für die Verwendung der IRT ist, dass hier das Partial-Credit-Modell zur Verfügung steht, das bei polytomen Items unterschiedlich große Schritte von Niveau zu Niveau (bspw. von null Punkten zu einem Punkt und von einem Punkt zu zwei Punkten) zulässt, innerhalb und zwischen Items. Darüber hinaus erlaubt das Partial-Credit-Modell die Nutzung von Items mit unterschiedlich vielen Stufen (bspw. dichotome und polytome Items; Bond und Fox 2012; Boone et al. 2014; Fox und Jones 1998).

Die ANOVAs wurden gemäß den Hypothesen mit a-priori Kontrasten durchgeführt, bei denen jeweils die Treatmentgruppen mit der Vergleichsgruppe verglichen wurden (*Simple*). Prinzipiell wurde dabei berücksichtigt, dass die einzelnen Schüler, von denen die Testdaten vorlagen, in Forscherteams arbeiteten, die in Klassen organisiert waren, die wiederum die Treatmentgruppe bildeten, daher wurden geschachtelte ANOVAs berechnet.

Da die Treatmentgruppen sich zwar nicht statistisch signifikant (*Wissenschaftliches Denken*: $H(3) = 3,40$, $p = 0,334$; *Methodenwissen*: $H(3) = 1,27$, $p = 0,735$; *Fachwissen*: $H(3) = 1,57$, $p = 0,665$), aber dennoch tendenziell in den Pre-Test-Resultaten unterschieden, wurden die Lernzuwächse als Residualgewinne berechnet und als solche für die Varianzanalysen verwendet (Walpuski und Sumfleth 2007). Die Residualgewinne stellen nicht die absoluten, sondern die relativen Lernzuwächse im Vergleich zur Ge-

samtstichprobe dar und können daher auch negative Werte annehmen.

Ergebnisse

In Abb. 7 sind die mittleren Residualgewinne der Gruppen für die Tests Wissenschaftliches Denken (A), Methodenwissen (B) und Fachwissen (C) dargestellt.

Es zeigt sich, dass die *Forschertipps*, die *Concept Cartoons* und die Kombination aus beiden in allen Bereichen zu höheren Lernzuwächsen führen als in der Vergleichsgruppe. Eine Ausnahme bilden die *Forschertipps* im Bereich Fachwissen, da hier die Treatmentgruppe hinter der Vergleichsgruppe zurückbleibt. Die Unterschiede im Lernzuwachs sind im Wissenschaftlichen Denken für die Gruppen mit *Forschertipps* und *Concept Cartoons* signifikant im Vergleich zur Vergleichsgruppe ($d = 0,47$ bzw. $0,43$). Die Ergebnisse der a-priori-Kontraste sind in Tab. 4 dargestellt.

Somit kann festgehalten werden, dass sich die Lernunterstützungen *Forschertipps* und *Concept Cartoons* positiv auf den Erwerb von Wissenschaftlichem Denken auswirken ($H_{\text{Wissenschaftliches Denken}}$). Für die Treatmentgruppe 3 (Doppel-treatment) trifft die Hypothese ebenfalls zu, die Unterschiede können jedoch nicht als signifikant ausgewiesen werden. Auch die Hypothesen $H_{\text{Methodenwissen}}$ und $H_{\text{Fachwissen}}$ können nur tendenziell bewertet werden, da die Unterschiede zur Vergleichsgruppe keine statistische Signifikanz aufweisen. Hier zeigt sich, dass alle drei Treatments (*Forschertipps*, *Concept Cartoons* sowie die Kombination aus beiden) im Vergleich zur Vergleichsgruppe tendenziell für den Erwerb von

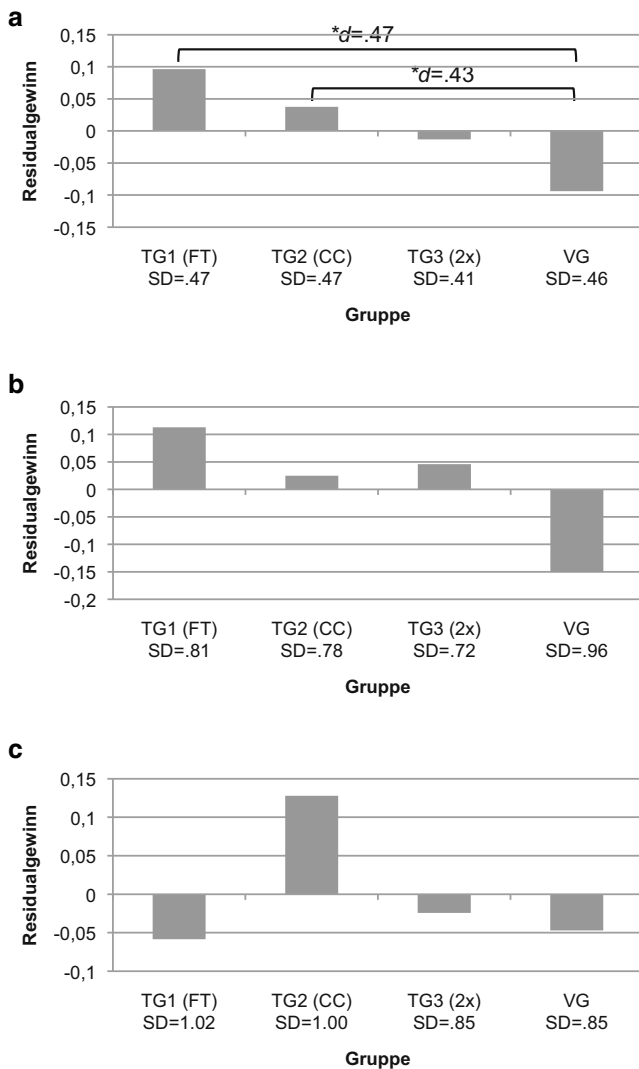


Abb. 7 Lernzuwächse (Residualgewinne) nach Gruppen – **a** Wissenschaftliches Denken, **b** Methodenwissen und **c** Fachwissen. TG1 (FT) Treatmentgruppe mit Forschertipps; TG2 (CC) Treatmentgruppe mit Concept Cartoons. TG3 (2x) Doppel-treatment; VG Vergleichsgruppe. *signifikanter Kontrast

Methodenwissen lernwirksam sind. Im Bereich Fachwissen zeigen sich die *Concept Cartoons* und das Doppel-treatment als tendenziell lernwirksam.

In Abb. 8 sind die Mittelwerte der Gruppen zum *Cognitive Load* angegeben. Es ist zu erkennen, dass die Vergleichsgruppe im Mittel die größte kognitive Belastung angibt.

Tab. 4 Residualgewinne – Wissenschaftliches Denken. A-priori-Kontraste

Kontrast	Quadratsumme	df	F	p (1-seitig)	d
VG vs. TG1 (Forschertipps)	0,88	1	4,93*	0,015	0,47
VG vs. TG2 (Concept Cartoons)	0,79	1	4,45*	0,019	0,43
VG vs. TG3 (Doppel-treatment)	0,19	1	1,08	0,151	/

*Signifikanter Kontrast

Die Ergebnisse der a-priori-Kontraste sind in Tab. 5 dargestellt. Es zeigt sich, dass alle Treatmentgruppen im Vergleich zur Vergleichsgruppe signifikant geringere kognitive Belastungen berichten. Diese Unterschiede gehen mit hohen Effektstärken einher ($0,72 < d < 1,40$), wobei der Effekt bei den *Concept Cartoons* am größten ist.

Somit kann die Hypothese $H_{\text{Cognitive Load}}$ für alle Treatmentgruppen bestätigt werden.

Diskussion und Desiderata

In der vorliegenden Studie wurden erstmals zwei Lernunterstützungs-Formate zur Förderung des prozeduralen Wissens sowie des Methodenwissens beim Forschenden Lernen in Bezug auf ihre Wirksamkeit für den Erwerb von Wissenschaftlichem Denken, Methodenwissen, Fachwissen und zur Reduktion der kognitiven Belastung vergleichend untersucht. Zentrale Befunde der Studie sind, dass die Lernunterstützung beim Forschenden Lernen (*Forschertipps*, *Concept Cartoons* sowie die Kombination aus beiden) zu einem höheren Lernertrag im Wissenschaftlichen Denken führen und zudem die kognitive Belastung beim Forschenden Lernen reduzieren können (Sweller et al. 1998). Für die Bereiche Methodenwissen und Fachwissen kann die Wirksamkeit nur tendenziell angenommen werden. Eine Ausnahme bildet die Gruppe mit *Forschertipps* im Bereich Fachwissen. Ähnliches konnte bereits von Hof (2011) festgestellt werden. Dort zeigte sich bereits, dass je nach Instruktion der Erwerb von wissenschaftsmethodischen Kompetenzen zu Lasten des Erwerbs von Fachwissen gehen kann.

Es ist darüber hinaus anzumerken, dass die *Forschertipps*, die auf die Förderung des prozeduralen Wissens abzielten (Atkinson et al. 2000), im Bereich Methodenwissen zu höheren Lernzuwächsen führten, als die *Concept Cartoons*, die zur expliziten Förderung des Methodenwissens konstruiert wurden (Sandoval und Reiser 2004). Zwar unterscheiden sich die Gruppen nicht signifikant in ihren Lernzuwächsen, aber hier sind weitere Untersuchungen notwendig, um die Wirkzusammenhänge zu analysieren. Bspw. ist es möglich, dass die *Forschertipps* ebenfalls zum reflektierten Handeln (*Warum* sind die einzelnen Schritte sinnvoll?) anregen und somit den Erwerb von Methodenwissen begünstigen (Keogh 1999).

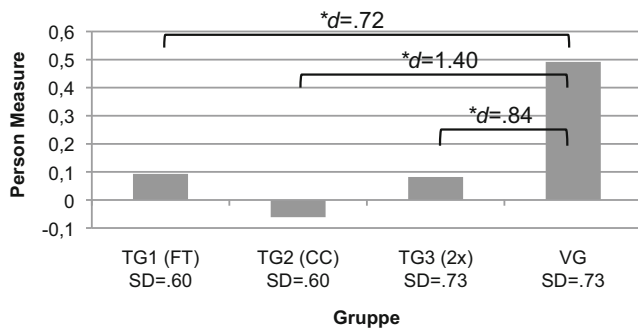


Abb. 8 Cognitive Load nach Gruppen

Zudem fällt auf, dass die Doppeltreatments hinter dem Einzel-treatment *Forschertipps* zurückbleibt. Hier kann die Auszählung der *Forschertipps* als Erklärung herangezogen werden. Es zeigte sich, dass im Doppeltreatment im Mittel 15 % und im Einzel-treatment ca. 22 % der *Forschertipps* genutzt wurden. Dies kann ein Indiz dafür sein, dass die Schüler im Doppeltreatment, auch wenn sie keine höhere kognitive Belastung berichten, aufgrund der Materialfülle die Lernunterstützungen nicht adäquat und vollständig nutzen konnten und daher geringere Lernzuwächse verzeichnen (Sweller et al. 1998), dies weist auf sog. Over-scripting hin (Dillenbourg 2002). Beim Doppeltreatment wurde mit der kompletten Anzahl Hilfen und Cartoons zu den Einzel-treatments additiv verfahren. Ob eine reduzierte Anzahl der Hilfsformate in der Kombination jedoch effektiver sein könnte, bleibt offen.

Zusammenfassend war das Forschende Lernen in dieser Konstellation effektiv, sogar für die Vergleichsgruppe. Dass die Gruppen sich nur wenig unterscheiden, (nur im wissenschaftlichen Denken statistisch signifikant), kann dadurch erklärt werden, dass alle Gruppen, schon relativ strukturierte Materialien (Experimentierhefte) mit Hinweisen auf die einzelnen zu berücksichtigenden Aspekte, die für die *Forschertipps* die Verlinkungen enthielten, hatten. In vergleichbaren Studien mit Kontrollgruppen, die normal, d. h. auch nicht im Stil des Forschenden Lernens, beschult wurden, sind ebenfalls nur kleine Effekte zu finden (Roesch et al. 2015). Darüber hinaus kann die Nutzung der Unterstützungen als Erklärung herangezogen werden. So wurden die *Forschertipps* nur relativ selten genutzt (durchschnittlich 17 %). Eine genauere Analyse der Bearbeitung der *Concept Cartoons* zeigt zudem, dass die Aussagen der Schüler

sich nur in ca. 64 % auf die Gütekriterien wissenschaftlichen Arbeitens bezogen und somit nur eingeschränkt zu den intendierten Diskussionen führten (Arnold et al. 2016). Hier sind weitere Untersuchungen notwendig, die die Nutzung der Lernunterstützungen optimieren. So sind zum Beispiel gezielte Trainings im Umgang mit den Unterstützungsformaten denkbar. Zukünftige Forschung sollte sich daher der weiteren Optimierung der Lernunterstützungen bzw. deren Nutzung widmen. Bspw. sollte überprüft werden, ob die Verklebung der *Forschertipps*, die der Kontrollierbarkeit dienen sollte, eine Hemmschwelle darstellt und sich negativ auf den Gebrauch auswirkt. Es ist zu erwarten, dass die *Forschertipps* häufiger genutzt werden, wenn diese frei zugänglich sind und somit auch lernwirksamer sein können. Für die *Concept Cartoons* ist denkbar, dass eine ausführlichere Schulung im Wissen um die Gütekriterien sich ggf. positiv auf die Diskussionen auswirken kann und somit die Lernförderlichkeit der *Concept Cartoons* gesteigert werden kann. Alternativ ist es denkbar, um die Qualität der Diskussionen zu fördern, diese im Plenum unter Moderation der Lehrkraft durchzuführen.

Nicht zuletzt konnte in anderen Studien gezeigt werden, dass die Leistungen in chemischen Experimentieraufgaben durch motivationale Faktoren, wie bspw. das Interesse an den zugrundeliegenden fachlichen Konzepten beeinflusst werden (Nehring et al. 2015). Dieser Einfluss kann für die Bearbeitung von Lernumgebungen und somit den Lernzuwachs wahrscheinlich noch umso stärker angenommen und als Erklärung für individuelle Varianzen und geringere Gruppenunterschiede herangezogen werden.

Schließlich kann an der Testgüte der Instrumente weitergearbeitet werden, wodurch verlässlichere und ggf. signifikante Ergebnisse zu erwarten wären, bspw. in Bezug auf die Reliabilitäten. Zudem weisen relativ hohe Out-Fit Werte auf Messfehler hin. Zukünftige Forschung sollte sich der Verbesserung dieser Instrumente widmen, um etwaige Unterschiede besser detektieren zu können. Darüber hinaus ist die relativ geringe Stichprobengröße der einzelnen Gruppen zu berücksichtigen. Generell ist anzumerken, dass die Studie in einer Jahrgangsstufe durchgeführt wurde, die Generalisierbarkeit der Ergebnisse auf andere Jahrgangsstufen sollte zukünftig sichergestellt werden.

Als Fazit lässt sich festhalten, dass Lernunterstützungen zur Förderung des prozeduralen Wissens sowie des Methodenwissens geeignet sind, den Erwerb Wissenschaftli-

Tab. 5 Cognitive Load. A-priori-Kontraste

Kontrast	Quadratsumme	df	F	p (1-seitig)	d
VG vs. TG1 (Forschertipps)	2,17	1	8,77*	0,002	0,72
VG vs. TG2 (Concept Cartoons)	8,28	1	33,39*	0,000	1,40
VG vs. TG3 (Doppeltreatment)	3,11	1	12,54*	0,001	0,84

*Signifikanter Kontrast

chen Denkens zu fördern (Mayer 2007). Mit dieser Studie wurden Fördermöglichkeiten von prozeduralem Wissen und deklarativem Methodenwissen in Form von jeweils einem Format überprüft. Es stellt sich die Frage, ob andere Formate ebenfalls oder sogar besser geeignet wären um diese Wissensarten zu fördern. Zukünftig sollten darüber hinaus etwaige Mediationseffekte untersucht werden und Kontrollvariablen, wie bspw. kognitive Fähigkeiten Berücksichtigung finden. Des Weiteren wäre es sinnvoll, die Wirkung der Lernunterstützungen differenziert nach Leistungsstärke zu untersuchen. Erste explorative Analysen zeigen, dass im Wissenschaftlichen Denken leistungsschwache Lernende besonders von Forschertipps und leistungsstarke Lernende stärker von den *Concept Cartoons* profitieren (Arnold 2015). Außerdem beschränkt sich diese Studie insgesamt auf die Untersuchung von *hard scaffolds*. Der Bereich der *soft scaffolds*, d. h. situativen Lehrerunterstützung, bedarf ebenfalls der weiteren Untersuchung. Diesbezüglich könnten die vorgestellten Formate auch eine Hilfestellung bieten.

Die gewonnenen Erkenntnisse über die Wirkung der Lernunterstützungen legen ferner nahe, dass das wissenschaftliche Denken durch weitere Variablen als das Methodenwissen, das Fachwissen und die allgemeinen kognitiven Fähigkeiten beeinflusst wird. Dies deckt sich mit den Befunden, dass durch diese Variablen 30 % der Varianzen im Wissenschaftlichen Denken erklärt werden können (Arnold 2015). Weitere mögliche Einflussfaktoren müssen zukünftig systematisch gesucht werden, um so das theoretische Modell zu erweitern und die Förderung zu optimieren.

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

- Abd-El-Khalick, F., & Lederman, N.G. (2000). Improving science teachers' conceptions of nature of science: a critical review of the literature. *International Journal of Science Education*, 22(7), 665–701.
- Abd-El-Khalick, F., BouJaoude, S., Duschl, R., Lederman, N.G., Mamlok-Naaman, R., Hofstein, A., Niaz, M., Treagust, D., & Tuan, H.I. (2004). Inquiry in science education: international perspectives. *Science Education*, 88(3), 397–419. doi:10.1002/sce.10118.
- Akerson, V.L., Abd-El-Khalick, F., & Lederman, N.G. (2000). Influence of a reflective explicit activity-based approach on elementary teachers' conceptions of nature of science. *Journal of Research in Science Teaching*, 37(4), 295–317.
- American Association for the Advancement of Science (AAAS) (2009). Benchmarks online. <http://www.project2061.org/publications/bsl/online/index.php?chapter=1#A0>. Zugegriffen: 02.12.2016.
- Arnold, J. (2015). *Die Wirksamkeit von Lernunterstützungen beim Forschenden Lernen: Eine Interventionsstudie zur Förderung des Wissenschaftlichen Denkens in der gymnasialen Oberstufe*. Berlin: Logos.
- Arnold, J., Kremer, K., & Mayer, J. (2013). Wissenschaftliches Denken beim Experimentieren – Kompetenzdiagnose in der Sekundarstufe II. In D. Krüger, A. Upmeyer zu Belzen, P. Schmiemann, A. Möller & D. Elster (Hrsg.), *Erkenntnisweg Biologiedidaktik 11* (S. 7–20). Kassel: Universitätsdruckerei.
- Arnold, J., Kremer, K., & Mayer, J. (2014). Understanding students' experiments – What kind of support do they need in inquiry tasks? *International Journal of Science Education*, 36(15–16), 2719–2749.
- Arnold, J., Kremer, K., & Mayer, J. (2016). Concept Cartoons als diskursiv-reflexive Szenarien zur Aktivierung des Methodenwissens beim Forschenden Lernen. *Biologie Lehren und Lernen. Zeitschrift für Didaktik der Biologie*, 20(1), 33–43.
- Atkinson, R.K., Sharon, J.D., Renkl, A., & Wortham, D. (2000). Learning from examples: instructional principles from the worked examples research. *Review of Educational Research*, 70(2), 181–214.
- Australian Curriculum Assessment and Reporting Authority (ACARA) (2012). Senior Secondary Curriculum – Biology. <http://www.australiancurriculum.edu.au/Static/docs/senior%20secondary/Senior%20Secondary%20Curriculum%20-%20Biology%20November%202012.pdf>. Zugegriffen: 02.12.2016.
- Beaumont-Walters, Y., & Soyibo, K. (2001). An analysis of high school students' performance on five integrated science process skills. *Research in Science & Technological Education*, 19(2), 133–145.
- Blanchard, M.R., Southerland, S.A., Osborne, J.W., Sampson, V.D., Annetta, L.A., & Granger, E.M. (2010). Is inquiry possible in light of accountability?: A quantitative comparison of the relative effectiveness of guided inquiry and verification laboratory instruction. *Science Education*, 94(4), 577–616.
- Bond, T.G., & Fox, C.M. (2012). *Applying the Rasch model – fundamental measurement in the human sciences* (2. Aufl.). New York: Routledge.
- Boone, W.J., & Scantlebury, K. (2006). The role of Rasch analysis when conducting science education research utilizing multiple-choice tests. *Science Education*, 90(2), 253–269.
- Boone, W.J., Staver, J.R., & Yale, M.S. (2014). *Rasch analysis in the human sciences*. Dordrecht: Springer.
- Buffler, A., Allie, S., Lubben, F., & Campbell, B. (2001). The development of first year physics students' ideas about measurement in terms of point and set paradigms. *International Journal of Science Education*, 23(11), 1137–1156.
- Bybee, R.W. (2006). Scientific inquiry and science teaching. In L.B. Flick & N.G. Lederman (Hrsg.), *Scientific inquiry and nature of science* (S. 1–14). Dordrecht: Springer.
- Chen, Z., & Klahr, D. (1999). All other things being equal: acquisition and transfer of the control of variables strategy. *Child Development*, 70(5), 1098–1120.
- Chinn, C.A., & Malhotra, B.A. (2002). Epistemologically authentic inquiry in schools: a theoretical framework for evaluating inquiry tasks. *Science Education*, 86(2), 175–218.
- Crawford, B.A. (2000). Embracing the essence of inquiry: New roles for science teachers. *Journal of Research in Science Teaching*, 37(9), 916–937.
- Crawford, B.A. (2007). Learning to teach science as inquiry in the rough and tumble of practice. *Journal of Research in Science Teaching*, 44(4), 613–642.
- Dean, D., & Kuhn, D. (2007). Direct instruction vs. discovery: the long view. *Science Education*, 91(3), 384–397.

- Department for Education and Skills/Qualification and Curriculum Authority (DfES/QCA) (2004). *Science – the national curriculum for England*. London: HMSO.
- Dillenbourg, P. (2002). Over-scripting CSCL: the risks of blending collaborative learning with instructional design. In P.A. Kirschner (Hrsg.), *Three worlds of CSCL. Can we support CSCL?* (S. 61–91). Heerlen: Open Universiteit Nederland.
- Dochy, F., Segers, M., Van den Bossche, P., & Gijbels, D. (2003). Effects of problem-based learning: a meta-analysis. *Learning and Instruction, 13*(5), 533–568.
- Duggan, S., & Gott, R. (2000a). Intermediate General National Vocational Qualification (GNVQ) science: a missed opportunity for a focus on procedural understanding? *Research in Science and Technological Education, 18*(2), 201–214.
- Duggan, S., Johnson, P., & Gott, R. (1996). A critical point in investigative work: defining variables. *Journal of Research in Science Teaching, 33*(5), 461–474.
- Duschl, R. (2000). Making the nature of science explicit. In R. Millar, J. Leech & J. Osborne (Hrsg.), *Improving science education: the contribution of research* (S. 187–206). Philadelphia: Open University Press.
- Forschergruppe Kassel (2007). Schritt für Schritt zur Lösung. Differenzierung durch Aufgaben mit gestuften Lernhilfen. *Naturwissenschaften im Unterricht – Physik, 18*(99/100), 42–45.
- Fox, C.M., & Jones, J.A. (1998). Uses of Rasch modeling in counseling psychology research. *Journal of Counseling Psychology, 45*(1), 30–45.
- Fraser, B.J. (1980). Development and validation of a test of enquiry skills. *Journal of Research in Science Teaching, 17*(1), 7–16.
- Furtak, E.M. (2006). The problem with answers: an exploration of guided scientific inquiry teaching. *Science Education, 90*(3), 453–467.
- Furtak, E.M., Seidel, T., Iverson, H., & Briggs, D.C. (2009). *Recent experimental studies of inquiry-based teaching: a meta-analysis and review*. Paper presented at the European Association for Research on Learning and Instruction, August 25–29, 2009, Amsterdam, Netherlands.
- Germann, P.J., & Aram, R.J. (1996a). Student performance on asking questions, identifying variables, and formulating hypotheses. *School Science and Mathematics, 4*, 192–201.
- Germann, P.J., & Aram, R.J. (1996b). Student performances on the science processes of recording data, analyzing data, drawing conclusions, and providing evidence. *Journal of Research in Science Teaching, 33*(7), 773–798.
- Germann, P.J., Aram, R., & Burke, G. (1996). Identifying patterns and relationships among the responses of seventh-grade students to the science process skill of designing experiments. *Journal of Research in Science Teaching, 33*(1), 79–99.
- Gijbels, D., Dochy, F., Van den Bossche, P., & Segers, M. (2005). Effects of problem-based learning: a meta-analysis from the angle of assessment. *Review of Educational Research, 75*(1), 27–61.
- Glaesser, J., Gott, R., Roberts, R., & Cooper, B. (2009a). The roles of substantive and procedural understanding in open-ended science investigations. *Research in Science Education, 39*(4), 595–624.
- Glaesser, J., Gott, R., Roberts, R., & Cooper, B. (2009b). Underlying success in open-ended investigations in science: using qualitative comparative analysis to identify necessary and sufficient conditions. *Research in Science & Technological Education, 27*(1), 5–30.
- Gott, R., & Duggan, S. (1995). *Investigative work in the science curriculum*. Buckingham: Open University Press.
- Gott, R., & Roberts, R. (2008). Concepts of evidence and their role in open-ended practical investigations and scientific literacy; background to published papers. https://www.dur.ac.uk/resources/education/research/res_rep_short_master_final.pdf. Zugegriffen: 02.12.2016.
- Gott, R., Duggan, S., & Roberts, R. (o. J.). Concepts of Evidence. Verfügbar unter: www.dur.ac.uk/rosalyn.roberts/Evidence/CofEv_Gott%20et%20al.pdf. Zugegriffen: 02.12.2016.
- Hammann, M., Phan, T.H., & Bayrhuber, H. (2007). Experimentieren als Problemlösen: Lässt sich das SDDS-Modell nutzen, um unterschiedliche Dimensionen beim Experimentieren zu messen? *Zeitschrift für Erziehungswissenschaft, 10*(8), 33–49.
- Hammann, M., Phan, T.H., Ehmer, M., & Grimm, T. (2008). Assessing pupils' skills in experimentation. *Journal of Biological Education, 42*(2), 66–72.
- Hänze, M., Schmidt-Weigand, F., & Stäudel, L. (2010). Gestuften Lernhilfen. In S. Boller & R. Lau (Hrsg.), *Individuelle Förderung durch Innere Differenzierung. Ein Praxishandbuch für Lehrerinnen und Lehrer der Sekundarstufe II* (S. 63–73). Weinheim: Beltz.
- Harlen, W. (1999). Purposes and procedures for assessing science process skills. *Assessment in Education: Principles, Policy & Practice, 6*(1), 129–144.
- Harms, U., Mayer, J., Hammann, M., Bayrhuber, H., & Kattmann, U. (2004). Kerncurriculum und Standards für den Biologieunterricht in der gymnasialen Oberstufe. In H.-E. Tenorth (Hrsg.), *Kerncurriculum Oberstufe II – Biologie, Chemie, Physik, Geschichte, Politik. Expertisen – im Auftrag der Ständigen Konferenz der Kultusminister (KMK)*. Weinheim: Beltz.
- Harwood, W.S. (2004). A new model for inquiry – is the scientific method dead? *Journal of College Science Teaching, 33*(7), 29–33.
- Hattie, J. (2009). *Visible learning: a synthesis of over 800 meta-analyses relating to achievement*. London: Routledge.
- Hmelo-Silver, C. (2004). Problem-based learning: What and how do students learn? *Educational Psychology Review, 16*(3), 235–266.
- Hmelo-Silver, C.E. (2006). Design principles for scaffolding technology-based inquiry. In A.M. O'Donnell, C.E. Hmelo-Silver & G. Erkens (Hrsg.), *Collaborative learning, reasoning, and technology* (S. 147–170). New York: Routledge.
- Hmelo-Silver, C.E., Duncan, R.G., & Chinn, C.A. (2007). Scaffolding and achievement in problem-based and inquiry learning: a response to Kirschner, Sweller, and Clark (2006). *Educational Psychologist, 42*(2), 99–107.
- Hof, S. (2011). *Wissenschaftsmethodischer Kompetenzerwerb durch Forschendes Lernen: Entwicklung und Evaluation einer Interventionsstudie*. Kassel: University Press.
- Hofstein, A., Navon, O., Kipnis, M., & Mamlok-Naaman, R. (2005). Developing students' ability to ask more and better questions resulting from inquiry-type chemistry laboratories. *Journal of Research in Science Teaching, 42*(7), 791–806.
- iSTAR Assessment (2011). Inquiry for scientific thinking and reasoning (vol. 2014). <http://www.istarassessment.org/>. Zugegriffen: 02.12.2016.
- de Jong, T., & van Joolingen, W.R. (1998). Scientific discovery learning with computer simulations of conceptual domains. *Review of Educational Research, 68*(2), 179–201.
- Keogh, B. (1999). Concept cartoons, teaching and learning in science: an evaluation. *International Journal of Science Education, 21*(4), 431–446.
- Khishfe, R., & Abd-El-Khalick, F. (2002). Influence of explicit and reflective versus implicit inquiry-oriented instruction on sixth graders' views of nature of science. *Journal of Research in Science Teaching, 39*(7), 551–578.
- Kirschner, P.A. (1992). Epistemology, practical work and academic skills in science education. *Science and Education, 1*(3), 273–299.
- Kirschner, P.A., Sweller, J., & Clark, R.E. (2006). Why minimal guidance during instruction does not work: an analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist, 41*(2), 75–86.
- Klahr, D., & Dunbar, K. (1988). Dual space search during scientific reasoning. *Cognitive Science, 12*(1), 1–48.
- Klahr, D., & Nigam, M. (2004). The equivalence of learning paths in early science instruction – effects of direct instruction and discovery learning. *Psychological Science, 15*(10), 661–667.

- Klahr, D., Fay, A.L., & Dunbar, K. (1993). Heuristics for scientific experimentation: a developmental study. *Cognitive Psychology*, 25(1), 111–146.
- KMK (2004). *Einheitliche Prüfungsanforderungen in der Abiturprüfung Biologie*. München: Luchterhand.
- Kremer, K., & Mayer, J. (2013). Entwicklung und Stabilität von Vorstellungen über die Natur der Naturwissenschaften. *Zeitschrift für Didaktik der Naturwissenschaften*, 19, 77–101.
- Kremer, K., Specht, C., Urhahne, D., & Mayer, J. (2013). The relationship in biology between the nature of science and scientific inquiry. *Journal of Biological Education*, 48(1), 1–8.
- Kuhn, D., & Dean, D. (2005). Is developing scientific thinking all about learning to control variables? *Psychological Science*, 16(11), 866.
- Kuhn, D., Black, J., Keselman, A., & Kaplan, D. (2000). The development of cognitive skills to support inquiry learning. *Cognition and Instruction*, 18(4), 495–523.
- Künsting, J. (2007). *Effekte von Zielqualität und Zielspezifität auf selbstreguliert-entdeckendes Lernen durch Experimentieren*. Essen: Universität Duisburg-Essen.
- Künsting, J., Wirth, J., & Paas, F. (2010). The goal specificity effect on strategy use and instructional efficiency during computer-based scientific discovery learning. *Comput. Educ.*, 56(3), 668–679.
- Kurth, B. (2015). Neue KMK-Präsidentin Kurth: „Das Zentralabitur ist eine Vision“, *SchulSpiegel*. <http://www.spiegel.de/schulspiegel/kmk-praesidentin-brunhild-kurth-das-zentralabitur-ist-eine-vision-a-1012892.html>. Zugegriffen: 02.12.2016.
- Landis, J.R., & Koch, G.G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159–174.
- Lederman, J.S., Lederman, N.G., Bartos, S.A., Bartels, S.L., Meyer, A.A., & Schwartz, R.S. (2014). Meaningful assessment of learners' understandings about scientific inquiry – the Views About Scientific Inquiry (VASI) questionnaire. *Journal of Research in Science Teaching*, 51(1), 65–83.
- Lederman, N.G., Abd-El-Khalick, F., Bell, R.L., & Schwartz, R.S. (2002). Views of nature of science questionnaire: toward valid and meaningful assessment of learners' conceptions of nature of science. *Journal of Research in Science Teaching*, 36(6), 497–521.
- Lienert, G., & Raatz, U. (1998). *Testaufbau und Testanalyse*. Weinheim: Beltz.
- Lin, X., & Lehman, J.D. (1999). Supporting learning of variable control in a computer-based biology environment: effects of prompting college students to reflect on their own thinking. *Journal of Research in Science Teaching*, 36(7), 837–858.
- Linacre, J.M. (2011). *Winsteps® Rasch measurement computer program user's guide*. Beaverton: Winsteps.com.
- Lubben, F., & Millar, R. (1996). Children's ideas about the reliability of experimental data. *International Journal of Science Education*, 18(8), 955–968.
- Lubben, F., Buffler, A., Campbell, B., & Allie, S. (2001). Point and set reasoning in practical science measurement by entering university freshmen. *Science Education*, 85(4), 311–327.
- Mayer, J. (2007). Erkenntnisgewinnung als wissenschaftliches Problemlösen. In D. Krüger & H. Vogt (Hrsg.), *Theorien in der biologiedidaktischen Forschung* (S. 177–186). Berlin: Springer.
- Mayer, J. (2013). Erkenntnisse mit naturwissenschaftlichen Methoden gewinnen. In H. Gropengießer, U. Harms & U. Kattmann (Hrsg.), *Fachdidaktik Biologie* (S. 56–61). Hallbergmoos: Aulis.
- Mayer, J., & Wellnitz, N. (2014). Die Entwicklung von Kompetenzstrukturmodellen. In D. Krüger, I. Parchmann & H. Schecker (Hrsg.), *Methoden in der naturwissenschaftsdidaktischen Forschung* (S. 19–29). Berlin: Springer.
- Mayer, J., & Ziemek, H.-P. (2006). Offenes Experimentieren. Forschendes Lernen im Biologieunterricht. *Unterricht Biologie*, 317, 4–12.
- Mayer, J., Harms, U., Hammann, M., Bayrhuber, H., & Kattmann, U. (2004). Kerncurriculum Biologie der gymnasialen Oberstufe. *Der Mathematische und Naturwissenschaftliche Unterricht*, 57(3), 166–173.
- Mayer, J., Grube, C., & Möller, A. (2008). Kompetenzmodell naturwissenschaftlicher Erkenntnisgewinnung. In U. Harms & A. Sandmann (Hrsg.), *Lehr- und Lernforschung in der Biologiedidaktik* (S. 63–79). Innsbruck: Studienverlag.
- Meier, M., & Mayer, J. (2011). Gewusst Vee! – Ein Diagnoseinstrument zur Erfassung von Konzept- und Methodenwissen im Biologieunterricht. *Schulpädagogik – heute*, 1(3), 1–12.
- National Research Council (NRC) (2012). *A framework for K-12 science education: practices, crosscutting concepts, and core ideas*. Washington, D.C.: National Academies Press.
- National Research Council (NRC) (1996). *National science education standards*. Washington, D.C.: National Academy Press.
- Naylor, S., & Keogh, B. (1999). Constructivism in classroom: theory into practice. *Journal of Science Teacher Education*, 10(2), 93–106.
- Nehring, A., Nowak, K.H., zu Belzen, A.U., & Tiemann, R. (2015). Predicting students' skills in the context of scientific inquiry with cognitive, motivational, and sociodemographic variables. *International Journal of Science Education*, 37(9), 1343–1363.
- Neumann, I. (2011). *Beyond physics content knowledge – modeling competence regarding nature of scientific inquiry and nature of scientific knowledge*. Berlin: Logos.
- Neumann, I., & Kremer, K. (2013). Nature of Science and epistemologische Überzeugungen – Ähnlichkeiten und Unterschiede. *Zeitschrift für Didaktik der Naturwissenschaften*, 19, 211–234.
- NGSS Lead States (2013). Next generation science standards: for states, by states. <http://www.nextgenscience.org/next-generation-science-standards>. Zugegriffen: 02.12.2016.
- Ontario Ministry of Education (OME) (2008a). *The Ontario curriculum, grades 9 and 10: science*. Toronto: Queen's Printer for Ontario.
- Ontario Ministry of Education (OME) (2008b). *The Ontario curriculum, grades 11 and 12: science*. Toronto: Queen's Printer for Ontario.
- Osborne, J., Collins, S., Ratcliffe, M., Millar, R., & Duschl, R. (2003). What “ideas-about-science” should be taught in school science? *Journal of Research in Science Teaching*, 40(7), 692–720.
- Özmen, H., Demircioglu, G., & Burhan, Y. (2012). Using laboratory activities enhanced with concept cartoons to support progression in students' understanding of acid-base concepts. *Asia-Pacific Forum on Science Learning and Teaching*, 13(1), Article 8.
- Phillips, K.A., & Germann, P.J. (2002). The inquiry “I”: a tool for learning scientific inquiry. *The American Biology Teacher*, 64(7), 512–520.
- Priemer, B. (2011). Was ist das Offene beim offenen Experimentieren? *Zeitschrift für Didaktik der Naturwissenschaften*, 17, 315–337.
- Puntambekar, S., & Hübscher, R. (2005). Tools for scaffolding students in a complex learning environment. *Educational Psychologist*, 40(1), 1–12.
- Renkl, A. (2005). The worked-out examples principle in multimedia learning. In R.E. Mayer (Hrsg.), *The Cambridge handbook of multimedia learning* (S. 229–245). New York: Cambridge University Press.
- Roberts, R. (2001). Procedural understanding in biology: the thinking behind the doing. *Journal of Biological Education*, 35(3), 113–117.
- Roberts, R., & Gott, R. (2003). Assessment of biology investigations. *Journal of Biological Education*, 37(3), 114–121.
- Roberts, R., & Gott, R. (2004). A written test for procedural understanding: a way forward for assessment in the UK science curriculum? *Research in Science & Technological Education*, 22(1), 5–21.
- Roesch, F., Nerb, J., & Riess, W. (2015). Promoting experimental Problemsolving ability in sixth-grade students through problem-oriented teaching of ecology: findings of an intervention study in a complex domain. *International Journal of Science Education*, 37(4), 577–598.

- Sandoval, W. A., & Morrison, K. (2003). High school students' ideas about theories and theory change after a biological inquiry unit. *Journal of Research in Science Teaching*, 40(4), 369–392.
- Sandoval, W. A., & Reiser, B. J. (2004). Explanation-driven inquiry: integrating conceptual and epistemic scaffolds for scientific inquiry. *Science Education*, 88(3), 345–372.
- Saye, J., & Brush, T. (2002). Scaffolding critical reasoning about history and social issues in multimedia-supported learning environments. *Educational Technology Research and Development*, 50(3), 77–96.
- Schmidt-Borcherding, F., Hänze, M., Wodzinski, R., & Rincke, K. (2013). Inquiring scaffolds in laboratory tasks: an instance of a “worked laboratory guide effect”? *European Journal of Psychology of Education*, 28(4), 1381–1395.
- Schmidt-Weigand, F., Franke-Braun, G., & Hänze, M. (2008). Erhöhen gestufte Lernhilfen die Effektivität von Lösungsbeispielen? *Unterrichtswissenschaft*, 36(4), 33–42.
- Schmidt-Weigand, F., Hänze, M., & Wodzinski, R. (2009). Complex problem solving and worked examples: the role of prompting strategic behavior and fading-in solution steps. *Zeitschrift für Pädagogische Psychologie*, 23(2), 129–138.
- Schwartz, R., Lederman, N., & Crawford, B. (2004). Developing views of nature of science in an authentic context: an explicit approach to bridging the gap between nature of science and scientific inquiry. *Science Education*, 88(4), 610–645.
- Schwartz, R. S., & Crawford, B. A. (2006). Authentic scientific inquiry as context for teaching nature of science. Identifying critical elements for success. In L. B. Flick & N. G. Lederman (Hrsg.), *Scientific inquiry and nature of science. Implications for teaching, learning and teacher education* (S. 331–355). Dordrecht: Kluwer.
- Schwartz, R. S., Lederman, N. G., & Lederman, J. S. (2008). *An instrument to assess views of scientific inquiry: the VOSI questionnaire*. International Conference of the National Association for Research in Science Teaching (NARST), Baltimore.
- Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (2005a). *Bildungsstandards im Fach Biologie für den Mittleren Schulabschluss*. München: Luchterhand.
- Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (2005b). *Beschlüsse der Kultusministerkonferenz – Bildungsstandards im Fach Chemie für den Mittleren Schulabschluss*. München: Luchterhand.
- Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (2005c). *Beschlüsse der Kultusministerkonferenz – Bildungsstandards im Fach Physik für den Mittleren Schulabschluss*. München: Luchterhand.
- Solano-Flores, G., Shavelson, R. J., & Bachman, M. (1999). On the development and evaluation of a shell for generating science performance assessments. *International Journal of Science Education*, 21(3), 293–315.
- Stanat, P., & Pant, H. A. (2012). Bildungsstandards für die Allgemeine Hochschulreife: Konzeption und Entwicklung. <http://www.iqb.hu-berlin.de/bista/abi>. Zugegriffen: 02.12.2016
- Sweller, J., van Merriënboer, J. G., & Paas, F. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, 10(3), 251–296.
- Tamir, P., Doran, R. L., & Oon Chye, Y. (1992). Practical skills testing in science. *Studies in Educational Evaluation*, 18(3), 263–275.
- Temiz, B. K., Tasar, M. F., & Tan, M. (2006). Development and validation of a multiple format test of science process skills. *International Education Journal*, 7(7), 1007–1027.
- Theyßen, H. (2014). Methodik von Vergleichsstudien zur Wirkung von Unterrichtsmedien. In D. Krüger, I. Parchmann & H. Schecker (Hrsg.), *Methoden in der naturwissenschaftsdidaktischen Forschung* (S. 67–79). Berlin: Springer.
- Tobin, K. G., & Capie, W. (1982). Development and validation of a group test of integrated science processes. *Journal of Research in Science Teaching*, 19(2), 133–141.
- Völzke, K., Arnold, J., & Kremer, K. (2013). Schüler planen und beurteilen ein Experiment – Denken und Verstehen beim naturwissenschaftlichen Problemlösen. *Zeitschrift für interpretative Schul- und Unterrichtsforschung (ZISU)*, 2(1), 58–86.
- Walpuski, M., & Sumfleth, E. (2007). Strukturierungshilfen und Feedback zur Unterstützung experimenteller Kleingruppenarbeit im Chemieunterricht. *Zeitschrift für Didaktik der Naturwissenschaften*, 13, 181–198.
- Wellnitz, N. (2012). *Kompetenzstruktur und -niveaus von Methoden naturwissenschaftlicher Erkenntnisgewinnung*. Berlin: Logos.
- Wellnitz, N., & Mayer, J. (2013). Erkenntnismethoden in der Biologie – Entwicklung und Evaluation eines Kompetenzmodells. *Zeitschrift für Didaktik der Naturwissenschaften*, 19, 315–345.
- Wichmann, A., & Leutner, D. (2009). Inquiry learning: multilevel support with respect to inquiry, explanations and regulation during an inquiry cycle. *Zeitschrift für Pädagogische Psychologie*, 23(2), 117–127.
- Wright, B. D., & Linacre, J. M. (1994). Reasonable mean-square fit values. *Rasch Measurement Transactions*, 8(3), 370.
- Zion, M., Slezak, M., Shapira, D., Link, E., Bashan, N., Brumer, M., Valanides, N., et al. (2004). Dynamic, open inquiry in biology learning. *Science Education*, 88(5), 728–753.