1 Titel

2 Welche Rückmeldungen wirken am besten?

3 Untertitel

4 Ergebnisse einer empirischen Untersuchung von Rückmeldeformen beim Planen von Experimenten

5 Zusammenfassung

6 Wichtiges Lernziel in den Naturwissenschaften sind Denk- und Arbeitsweisen der Erkenntnisgewinnung. Diese
7 Denk- und Arbeitsweisen gilt es im naturwissenschaftlichen Unterricht gezielt zu adressieren, damit
8 Schülerinnen und Schüler eine naturwissenschaftliche Grundbildung erlangen können. Im Prozess der
9 Erkenntnisgewinnung nimmt das Experimentieren als eine Denk- und Arbeitsweise eine wichtige Rolle ein, da
10 anhand der aus einem Experiment gewonnenen Daten neue Erkenntnisse generiert werden können. Um anhand
11 der gewonnenen Daten eindeutige Aussagen treffen zu können, muss die Variablenkontrollstrategie bereits bei
12 der Planung des Experiments angewendet werden. Schülerinnen und Schüler haben jedoch Schwierigkeiten die
13 Strategie anzuwenden und benötigen bei der Planung von Experimenten Unterstützung. Eine mögliche
14 Unterstützungsmaßnahme stellen lernbegleitende Rückmeldungen dar. Im Hinblick auf die Wirkung von
15 Rückmeldeformen auf das Planen von Experimenten liegen widersprüchliche Befunde vor. Ziel des vorliegenden
16 Beitrags ist daher die Entwicklung und Evaluation von Rückmeldeformen als Unterstützungsmaßnahme zur
17 Förderung der Anwendung der Variablenkontrollstrategie beim Planen von Experimenten. Die Maßnahme
18 wurde anhand einer Interventionsstudie im Fach Chemie evaluiert. Dabei wurden lösungsprozessbezogene
19 Rückmeldungen, lernstandbezogene Rückmeldungen und die Selbsteinschätzung als Rückmeldung kontrastiert.
20 Die Rückmeldeformen unterscheiden sich hinsichtlich der Informationen zum Lernziel, zum Lernstand und zu
21 den nächsten Lernschritten. Die Ergebnisse aus der Interventionsstudie belegen, dass die Rückmeldeform, die
22 Informationen zum Lernziel, zum Lernstand und zu nächsten Lernschritten liefert, zum höchsten Lernertrag
23 führt. In der Diskussion des Beitrags werden die Befunde in den derzeitigen Forschungsstand eingeordnet und
24 Perspektiven für weitere Forschung aufgezeigt.

25 Schlüsselwörter

26 Forschendes Lernen, Variablenkontrollstrategie, Rückmeldungen, Interventionsstudie

28 Title

29 Which type of feedback works best?

30 Subtitle

31 Analyzing the effect of different feedback types on students experimental designs

32 Summary

33 Scientific inquiry is an important approach for teaching and learning science. By doing inquiry new knowledge
34 can be generated. However, students' have to learn how to do inquiry. To plan an investigation is one important
35 practice when doing inquiry. Only based on the results of a controlled investigation verified conclusions can be
36 drawn. Thus, students have to apply the control-of-variable-strategy to their experimental plans. However,
37 several difficulties were reported in previous studies that highlight the importance of giving students individual
38 support during inquiry phases. One possible approach to support students' is formative assessment and especially
39 feedback. Its effect on students' learning was subject of several studies that report inconsistent results. The aim
40 of the presented work is the development and evaluation of different types of feedback that support the
41 application of the control-of-variable-strategy when planning investigations. The effectiveness of different
42 feedback types was evaluated by an intervention study in the field of chemistry education. Feedback that
43 provides information about the learning goal, the current learning level and possible next steps in learning was
44 compared with feedback that only provides information about the learning goal and the current learning level as
45 well as with self-assessment. The results indicate that feedback focusing on the learning goal, the current
46 learning level and next steps in learning is superior in comparison to the other two feedback types. In the
47 discussion, the present results are related to the results of previous studies in this research field. Besides
48 perspectives for further research are highlighted.

Key words

Scientific inquiry, control-of-variable-strategy, feedback, intervention study

Einleitung

Der Prozess der naturwissenschaftlichen Erkenntnisgewinnung wird als Ausführung von Denk- und Arbeitsweisen beschrieben, durch die neue Erkenntnisse über naturwissenschaftliche Konzepte generiert werden sollen (Abd-el-Khalick et al., 2004). Neun Denk- und Arbeitsweisen konnten theoretisch anhand einer Literaturrecherche identifiziert werden (Rönnebeck, Bernholt & Ropohl, 2016). Diese sind konstituierend für den Prozess der Erkenntnisgewinnung und lauten: 1. Forschungsfragen identifizieren, 2. Informationen recherchieren, 3. Hypothesen formulieren, 4. Experimente planen und durchführen, 5. Daten analysieren und interpretieren, 6. Erklärungen entwickeln, 7. Modelle konstruieren, 8. Argumentieren und 9. Kommunizieren.

Im naturwissenschaftlichen Unterricht gilt es nicht nur fachliche Konzepte, sondern auch die Denk- und Arbeitsweisen des Erkenntnisgewinnungsprozesses und ihre Anwendung gezielt zu fördern (Abd-el-Khalick et al., 2004; KMK, 2005; NRC, 2012). Forschendes Lernen ist eine Unterrichtskonzeption, die den Prozess der Erkenntnisgewinnung bzw. einzelne Denk- und Arbeitsweisen in den Mittelpunkt des Lehrens und Lernens stellt (Höttecke, 2010; Ropohl, Rönnebeck & Scheuermann, 2015). Schülerinnen und Schüler führen gezielt naturwissenschaftliche Denk- und Arbeitsweisen durch und bauen dabei prozessbezogene Kompetenzen auf. Eine im Kontext des forschenden Lernens wichtige Denk- und Arbeitsweise ist das Experimentieren. Anhand der mit einem Experiment gewonnenen Daten können Schlussfolgerungen hinsichtlich zuvor aufgestellter Fragestellungen und Hypothesen zu Ursache-Wirkungs-Beziehungen gezogen werden (Schulz, Staraschek & Wirtz, 2012; Wellnitz & Mayer, 2013). Das Experimentieren umfasst neben der Durchführung eines Experiments auch die vorherige Planung des Experiments unter Anwendung der Variablenkontrollstrategie (Vorholzer, Aufschneider & Kirschner, 2016). Bisherige empirische Befunde belegen, dass Schülerinnen und Schüler Unterstützung beim Planen von Experimenten benötigen (Arnold, Kremer & Mayer, 2017; Wahser & Sumfleth, 2008; Wellnitz & Mayer, 2013). In diesem Beitrag wird die Evaluation von Rückmeldungen als Unterstützungsmaßnahme bei der Planung von Experimenten dargestellt.

Anwendung der Variablenkontrollstrategie beim Experimentieren

Das Experimentieren als naturwissenschaftliche Denk- und Arbeitsweise umfasst aus wissenschaftstheoretischer Perspektive unterschiedliche Strategien zur Gewinnung von Erkenntnissen (Höttecke & Rieß, 2015; Schulz, Staraschek & Wirtz, 2015). Diesen Strategien ist gemein, dass sie „als Bemühen um Kongruenz experimenteller Praktiken mit theoretischen Annahmen“ (Höttecke & Rieß, 2015, S. 136) gedeutet werden können. Aus naturwissenschaftsdidaktischer Perspektive wird im Unterricht häufig eine starke Engführung vorgenommen, indem auf Aspekte wie Wiederholbarkeit oder Variablenkontrolle fokussiert wird. Höttecke und Rieß (2015) kritisieren, dass dadurch kein authentisches Bild von Naturwissenschaften aufgebaut werden kann. Aus wissenschaftstheoretischer Perspektive gilt es zu betonen, dass die Variablenkontrollstrategie eine Strategie beim Experimentieren ist, aber nicht die einzige.

Die Variablenkontrollstrategie muss bei der Planung eines Experiments angewendet werden, wenn zuvor aufgestellte Hypothesen zu Ursache-Wirkungs-Beziehungen empirisch geprüft werden sollen (Chen & Klahr, 1999; Wellnitz & Mayer, 2013). Bei richtiger Anwendung dieser Strategie wird eine Variable als unabhängige Variable in ihren Ausprägungen variiert, um den Einfluss dieser unabhängigen Variable auf eine abhängige Variable zu untersuchen. Die Auswahl der unabhängigen Variable richtet sich nach der Hypothese und dem darin ausgedrückten vermuteten Zusammenhang. Alle übrigen möglichen Einflussvariablen müssen als Störvariablen konstant gehalten werden. Auf diese Weise wird der Einfluss unter konstanten Bedingungen gemessen. Empirische Befunde belegen, dass Schülerinnen und Schüler Schwierigkeiten haben, die Variablenkontrollstrategie beim Planen von Experimenten adäquat anzuwenden (Arnold et al., 2017; Bullock & Sodian, 2003; Hammann, Phan, Ehmer & Grimm, 2008). Basierend auf diesen Befunden kann das Variieren der unabhängigen Variable unter kontrollierten Bedingungen insbesondere in offenen Aufgaben wie etwa einer Experimentplanung als Schwierigkeit der Schülerinnen und Schüler benannt werden. Von Schülerinnen und Schülern werden vorwiegend Experimente geplant, in denen eine zur Hypothese passende unabhängige Variable variiert wird, jedoch keine kontrollierten Bedingungen durch Konstanthalten der Störvariablen sichergestellt

werden. Der systematische Umgang mit Variablen stellt eine weitere Schwierigkeit dar. Schülerinnen und Schüler variieren Ausprägungen der falschen unabhängigen Variable, also einer Variable, deren Variation keine Überprüfung der Hypothese ermöglicht.

Aufgrund der zentralen Stellung des Experiments im Prozess der Erkenntnisgewinnung und aufgrund der berichteten Schwierigkeiten seitens der Schülerinnen und Schüler bei der Planung von Experimenten unter kontrollierten Bedingungen kann geschlussfolgert werden, dass es Maßnahmen zu entwickeln und zu evaluieren gilt, die Schülerinnen und Schüler in der Anwendung der Variablenkontrollstrategie unterstützen.

Rückmeldungen als Unterstützungsmaßnahme

Eine potenziell effektive Möglichkeit, Schülerinnen und Schüler beim Planen von Experimenten und bei der Anwendung der Variablenkontrollstrategie zu unterstützen, stellen Rückmeldungen dar (Hattie & Timperley, 2007; Ross, 1988; Schwichow, Zimmerman, Croker & Härtig, 2016). In Meta-Analysen haben sich Rückmeldungen im Hinblick auf den Lernzuwachs der Schülerinnen und Schüler generell als eine sehr wirksame Unterstützungsmaßnahme erwiesen ($d = .26$, Bangert-Drowns, Kulik, Kulik & Morgan, 1991; $d = .73$, Hattie, 2009; $d = .46$, Kluger & DeNisi, 1996). Allerdings hat nicht jede Form von Rückmeldung einen positiven Effekt auf das Lernen. Hohe positive Effekte haben beispielsweise Rückmeldungen, die Schülerinnen und Schülern aufzeigen, inwiefern sie sich zu ihrem vorherigen Lernstand verbessert haben ($d = .55$, Kluger & DeNisi, 1996). Ferner sind Rückmeldungen lernwirksamer, wenn sie sich auf vorher festgesetzte Lernziele beziehen ($d = .51$, Kluger & DeNisi, 1996).

Basierend auf diesen und weiteren Befunden schlussfolgern Hattie und Timperley (2007) drei inhaltliche Elemente lernwirksamer Rückmeldungen: (1) Rückmeldungen sollen Lernenden aufzeigen, was das Lernziel ist und welche Kriterien das Erreichen des Lernziels kennzeichnen, (2) sie sollen die Schülerinnen und Schüler über den aktuellen Lernstand informieren und (3) sie sollen den Schülerinnen und Schülern nächste Lernschritte vorschlagen, die zum Erreichen des Lernziels durchgeführt werden müssen. Lernwirksame Rückmeldungen enthalten damit Informationen zum Lernziel, Lernstand und nächsten Lernschritten.

Wenn Rückmeldungen genutzt werden sollen, um einen Lernprozess stärker auf ein Lernziel hin auszurichten, so muss die Diagnose und Bewertung lernbegleitend erfolgen. In der englischsprachigen Literatur hat sich für dieses Prinzip der Unterstützung der Begriff *formative assessment* (dt.: formative Diagnose und Bewertung) etabliert (Black & Wiliam, 1998). Rückmeldungen gelten in der Konzeption der formativen Diagnose und Bewertung als eine wichtige Komponente, die Transparenz des Lernziels als eine Weitere. Hinsichtlich der Lernwirksamkeit der Konzeption ist die Befundlage nicht eindeutig und schwer vergleichbar, da in bisherigen Analysen die zugrundeliegende Definition bzw. die Rahmenbedingungen von Interventionen nur unzureichend beschrieben wurden. Ergebnisse einer kriteriengeleiteten Meta-Analyse berichten Kingston und Nash (2011), wonach formative Diagnose und Bewertung in den Naturwissenschaften einen kleinen Effekt hat ($d = .09$, 95 % $CI = -.09$ bis $.25$).

Empirische Befunde zur Wirksamkeit von Rückmeldungen liegen zudem speziell für die Anwendung der Variablenkontrollstrategie vor. Schwichow et al. (2016) berichten basierend auf einer Meta-Analyse keine signifikante Überlegenheit von Interventionen mit lernbegleitenden Rückmeldungen ($g = 0.66$) im Vergleich zu Interventionen ohne Rückmeldungen ($g = 0.58$). Demgegenüber berichtet Ross (1988) basierend auf einer Meta-Analyse sehr wohl eine Überlegenheit von Interventionen mit lernbegleitenden Rückmeldungen und einen hohen Effekt auf die Anwendung der Variablenkontrollstrategie ($d = 1.20$). Diese widersprüchlichen Befunde legen weitere Forschungsarbeiten in diesem Feld nahe. Es gibt Beispiele von Interventionsstudien, in denen die Wirkung unterschiedlicher Rückmeldeformen auf die Anwendung der Variablenkontrollstrategie untersucht wurden (Vollmeyer & Rheinberg, 2005; Wollenschläger, Möller & Harms, 2012). Diesen Studien ist gemein, dass sie die Bedingungen, unter denen Rückmeldungen lernwirksam sind, untersuchen. In der Untersuchung von Wollenschläger et al. (2012) haben sich Rückmeldungen als lernförderlich erwiesen, wenn sie Informationen zum Lernziel, Lernstand und nächsten Lernschritten (lösungsprozessbezogene Rückmeldungen) enthalten und wenn sie von der Testleitung gegeben werden. Ein weiterer Effekt lösungsprozessbezogener Rückmeldungen durch die Testleitung ist eine höhere Wahrnehmung der Unterstützung auf Seiten der Schülerinnen und Schüler. Keinen Lernzuwachs konnten Schülerinnen und Schüler verzeichnen, die ihren Lernstand selbst einschätzen mussten und die keine Rückmeldungen erhalten haben (Vollmeyer & Rheinberg, 2005). Doch auch

Rückmeldungen, die nur den Lernstand zurückmelden (lernstandbezogene Rückmeldungen), können lernwirksam sein.

Bei einer Selbsteinschätzung schätzen Schülerinnen und Schüler ihren eigenen Lernstand ein und schlussfolgern eigenständig nächste Lernschritte zur Erreichung des vorgegebenen Lernziels (McMillan & Hearn, 2008). Entscheidend für die Selbsteinschätzung ist, dass die Kriterien zur Erreichung des Lernziels eindeutig definiert sind (Adams & King, 1995). Befunde einer Meta-Analyse zeigen, dass Selbsteinschätzungen den Lernprozess unterstützen können ($d = .47$, Falchikov & Boud, 1989; Falchikov & Goldfinch, 2000), gerade auch im Kontext des naturwissenschaftlichen Unterrichts (White & Frederiksen, 1998). Es wird vermutet, dass Schülerinnen und Schüler aufgrund mehrerer Bedingungen unterstützt werden, zum Beispiel durch die intensive Auseinandersetzung mit den Kriterien zur Erreichung des Lernziels (Andrade & Du, 2007). Doch auch die Befundlage zum Self-Assessment (dt.: Selbsteinschätzung) ist nicht konsistent.

In Gänze zeigen die Ergebnisse bisheriger Untersuchungen kein einheitliches Bild zu den Bedingungen, unter denen Rückmeldungen die Anwendung der Variablenkontrollstrategie am besten fördern. Um Aussagen darüber treffen zu können, sind weitere Untersuchungen notwendig. Beachtet werden muss dabei, dass das Lernziel stets gezielt eingeführt werden sollte und dass der Effekt der Rückmeldungen auf den Lernzuwachs vermutlich über die wahrgenommene Unterstützung vermittelt wird.

Ziel sowie Fragestellungen und Hypothesen der Untersuchung

Vor dem Hintergrund des bisherigen Forschungsstandes ist das Ziel der nachfolgend dargestellten Untersuchung die empirische Überprüfung der Lernwirksamkeit lernbegleitender Rückmeldeformen beim forschenden Lernen zur Förderung der Anwendung der Variablenkontrollstrategie beim Planen von Experimenten. Es werden zwei Forschungsfragen und basierend auf dem berichteten Forschungsstand Hypothesen formuliert. Aufgrund der in Teilen widersprüchlichen Befundlage wurde Hypothese 1b als Alternativhypothese zu Hypothese 1a aufgestellt.

Forschungsfrage 1: Welche Rückmeldeform fördert das Planen von Experimenten unter Anwendung der Variablenkontrollstrategie am besten?

Hypothese 1a: Lösungsprozessbezogene Rückmeldungen mit Informationen zu Lernziel, Lernstand und nächsten Lernschritten unterstützen das Planen von Experimenten am besten (Hattie & Timperley, 2007; Wollenschläger et al., 2012).

Hypothese 1b: Lernstandbezogene Rückmeldungen mit den Informationen zu Lernziel und Lernstand unterstützen Schülerinnen und Schüler gleichermaßen wie lösungsprozessbezogene Rückmeldungen mit Informationen zu Lernziel, Lernstand und nächsten Lernschritten (Vollmeyer & Rheinberg, 2005).

Hypothese 1c: Rückmeldungen in Form von Selbsteinschätzungen, bei denen die Schülerinnen und Schüler anhand des Lernziels ihren Lernstand und ihre nächsten Lernschritte selbst einschätzen, bewirken den geringsten Lernzuwachs im Planen von Experimenten (Wollenschläger et al., 2012).

Forschungsfrage 2: Inwiefern wird der Effekt der Rückmeldungen auf den Lernzuwachs im Planen von Experimenten über die wahrgenommene Unterstützung vermittelt?

Hypothese 2: Die wahrgenommene Unterstützung der Rückmeldungen weist einen direkten Effekt auf den Lernzuwachs beim Planen von Experimenten auf. Zudem mediiert die wahrgenommene Unterstützung den Effekt der Rückmeldungen auf den Lernzuwachs beim Planen von Experimenten (Wollenschläger et al., 2012).

Design und Methode der Untersuchung

Zur Beantwortung der beiden Forschungsfragen wurde eine quasi-experimentelle Interventionsstudie im Prä-Post-Follow up-Testdesign mit drei Interventionsgruppen durchgeführt. Nachfolgend werden die unabhängigen und abhängigen Variablen, das Interventionsdesign, der Interventionsgegenstand, der Treatment-Check, die Kontroll- und Hintergrundvariablen, die Stichprobe und die Datenanalyse beschrieben.

unabhängige Variablen: Rückmeldeformen

Als unabhängige Variable wird in der Intervention die Rückmeldeform variiert. Insgesamt gibt es drei Ausprägungen: die lösungsprozessbezogene Rückmeldung, die lernstandbezogene Rückmeldung und die Selbsteinschätzung (Tab. 1). Während die Schülerinnen und Schüler der Interventionsgruppen 1 und 2 eine individuelle Rückmeldung erhalten, schätzen die Schülerinnen und Schüler der Gruppe 3 ihren Lernstand und nächste Lernschritte durch einen Vergleich ihrer Planung mit dem Lernziel selbst ein. Um die Vergleichbarkeit der Rückmeldungen zu gewährleisten, wurden entsprechend des Lernstandes standardisierte Formulierungen für die Erstellung gewählt.

Tab. 1: Rückmeldeformen in den Interventionsgruppen

Informationen der Rückmeldung	Interventionsgruppen		
	1 (Lösungsprozessbezogene Rückmeldung)	2 (Lernstandbezogene Rückmeldung)	3 (Selbsteinschätzung)
Lernziel	gegeben	gegeben	gegeben
Lernstand	gegeben	gegeben	offen
nächste Lernschritte	gegeben	offen	offen

abhängige Variable: Planung von Experimenten

Die abhängige Variable ist die Kompetenz der Schülerinnen und Schüler ein Experiment unter Anwendung der Variablenkontrollstrategie zu planen. Es wird erwartet, dass die drei Rückmeldeformen unterschiedlich gut auf diese Kompetenz nach Einführung des Lernziels wirken. Die Lernwirksamkeit wird anhand des Lernzuwachses im Wissen zu naturwissenschaftlichen Arbeitsweisen und in der Kompetenz zur Planung von Experimenten gemessen.

Das Wissen zu naturwissenschaftlichen Arbeitsweisen wurde in der Intervention implizit gefördert und stellt daher eine zur Intervention distale Variable dar. Es wurde zum Prä-, Post- und Follow Up-Messzeitpunkt anhand des Naturwissenschaftliche Arbeitsweisen-Tests (adaptiert nach Klos, 2007) erhoben. Als Testzeit standen jeweils 20 Minuten zur Verfügung. Der Test weist zu allen drei Messzeitpunkten eine gute Reliabilität auf ($.79 \leq \alpha \leq .82$).

Das Planen von Experimenten unter Anwendung der Variablenkontrollstrategie ist eine zur Intervention proximale Variable, da das Planen von Experimenten explizit in der Intervention gefördert wurde. Die Fähigkeit des Planens von Experimenten wurde anhand einer offenen Aufgabe in 10 Minuten (adaptiert nach Klos, 2007; s. online Ergänzung) zum Prä-, Post- und Follow Up-Messzeitpunkt gemessen. Für jede Experimentplanung wurde das Erreichen von zehn Merkmalen auf einer zweistufigen Skala überprüft (Tab. 2; s. online Ergänzung für vollständiges Bewertungsschema). Für 10 % der Experimentplanungen zum Post-Messzeitpunkt wurde eine Doppelkodierung zur Bestimmung der Interraterreliabilität vorgenommen. Diese liegt in einem guten bis sehr guten Bereich ($.70 \leq \tau \leq 1.0$). Die Reliabilität des Instruments liegt zu den drei Messzeitpunkten in einem für Gruppenvergleiche ausreichend hohen Bereich ($.50 \leq \alpha \leq .69$; Lienert & Raatz, 1998). Für die vertiefende Analyse der Qualität der Experimentplanungen wurden fünf Qualitätsstufen aus der Literatur abgeleitet (Tab. 3; u.a. adaptiert nach Bullock & Sodian, 2003; Hammann et al., 2006; Siler & Klahr, 2012) und die Experimentplanungen diesen Stufen zugeordnet.

Tab. 2: Bewertungsschema für eines von zehn Merkmalen einer Experimentplanung

Merkmal der Experimentplanung	Stufe 0	Stufe 1	Stufe 2
Variation der unabhängigen Variable unter Angabe einer Messdauer	Die unabhängige Variable wird nicht variiert.	Die unabhängige Variable wird ohne Angabe einer Messdauer variiert.	Die unabhängige Variable wird unter Angabe einer Messdauer variiert.

Tab. 3: Qualitätsstufen der Experimentplanungen hinsichtlich der Anwendung der Variablenkontrollstrategie

Qualität	Beschreibung
5	Das Experiment beinhaltet eine Variation der unabhängigen Variable unter stark kontrollierten Bedingungen (Messdauer und mehr als zwei Störvariablen).

- 4 Das Experiment beinhaltet eine Variation der unabhängigen Variable unter kontrollierten Bedingungen (Messdauer und ein oder zwei Störvariablen).
- 3 Das Experiment beinhaltet eine Variation der unabhängigen Variable bei festgelegter Messdauer.
- 2 Das Experiment beinhaltet eine unkontrollierte Variation der unabhängigen Variable.
- 1 Das Experiment beinhaltet keine Variation der unabhängigen Variable.

1 **Interventionsdesign und -gegenstand**

2 Zur Untersuchung der Lernwirksamkeit der drei Rückmeldeformen wurden diese in drei Interventionsgruppen
3 kontrastiert (Tab. 1). Die drei Rückmeldeformen wurden lernbegleitend im Sinne der Konzeption formativer
4 Diagnose und Bewertung realisiert. Daher entspricht der zeitliche Verlauf der Intervention dem theoretischen
5 Verlauf der Konzeption (Dolin, Black, Harlen & Tiberghien, 2018). Die Intervention umfasste vier
6 Unterrichtsstunden. In der ersten Interventionsstunde wurden die Merkmale der Experimentplanung und in
7 diesem Zusammenhang vor allem die Variablenkontrollstrategie als Lernziel eingeführt. In Anlehnung an das
8 geführte forschende Lernen haben die Schülerinnen und Schüler in den nachfolgenden Interventionsstunden ein
9 Experiment auf sogenannten Forscherbögen zu einer vorgegebenen Forschungsfrage und einer vorgegebenen
10 Hypothese schriftlich geplant (Abb. 1). Anschließend wurde bewertet, welche Merkmale der
11 Variablenkontrollstrategie bei der Experimentplanung angewendet wurden. Basierend auf der Bewertung wurde
12 eine individuelle Rückmeldung formuliert, die jeder Schülerin bzw. jedem Schüler zu Beginn der nächsten
13 Stunde auf einem Rückmeldebogen gegeben wurde. Auf einem Reflexionsbogen konnte sie bzw. er ihren bzw.
14 seinen aktuellen Lernstand reflektieren. Insgesamt haben die Schülerinnen und Schüler drei Experimente geplant
15 und dementsprechend dreimal eine Rückmeldung zu ihrer Planung erhalten.

16 Das Thema ‚Metalle und ihre Bedeutung für Zivilisation und Umwelt‘ ist laut Lehrplan für den
17 Anfangsunterricht vorgesehen (Ministerium für Bildung, Wissenschaft, Forschung und Kultur des Landes
18 Schleswig-Holstein, 1997) und wurde zum Zeitpunkt der Intervention regulär unterrichtet. Als Themenaspekte
19 wurden für die Intervention die Eigenschaften von Metallen sowie die Reaktion von Metallen mit Sauerstoff
20 gewählt, da diese sowohl die Vorgaben der nationalen Bildungsstandards (KMK, 2005) als auch der
21 Kernlehrpläne anderer Bundesländer (Ministerium für Schule und Weiterbildung des Landes Nordrhein-
22 Westfalen, 2014) aufgreifen. Die Themenaspekte können als etablierte Unterrichtsinhalte angesehen werden.
23 Insbesondere zum zweiten Themenaspekt sind fachlich nicht angemessene Schülervorstellungen dokumentiert
24 (Hadfield, 1995), die in Form von Hypothesen Ausgangspunkte für Experimentplanungen sein können (Ropohl,
25 Scheuermann & Rönnebeck, 2015; Zenner, Scheuermann & Ropohl, 2017).

Prä-Messzeitpunkt: Kompetenz zum Planen von Experimenten (10'), Wissen zu den naturwissenschaftlichen Arbeitsweisen (20'), Fachwissen zum Thema Metalle (15'), kognitive Fähigkeiten (12')	Intervention				Reflexion der Rückmeldung Post-Messzeitpunkt: Kompetenz zum Planen von Experimenten (10'), Wissen zu den naturwissenschaftlichen Arbeitsweisen (20'), Fachwissen zum Thema Metalle (15'), wahrgenommene Unterstützung (5')	← 10 Wochen →	Follow up-Messzeitpunkt: Kompetenz zum Planen von Experimenten (10'), Wissen zu den naturwissenschaftlichen Arbeitsweisen (20'), Fachwissen zum Thema Metalle (15')
	Erste Interventionsstunde	Zweite Interventionsstunde	Dritte Interventionsstunde	Vierte Interventionsstunde			
	Einführung des Lernziels, hier der Anwendung der Variablenkontrollstrategie	Planung eines Experiments	Reflexion der Rückmeldung Planung eines Experiments	Reflexion der Rückmeldung Planung eines Experiments			
		Vorgegebene Hypothese: „Kupfer weist eine höhere Wärmeleitfähigkeit als Eisen und Aluminium auf.“	Vorgegebene Hypothese: „Bei dem schwarzen Belag handelt es sich um aus der Flamme stammenden Ruß.“	Vorgegebene Hypothese: „Der schwarze Belag entsteht bei der Reaktion von Kupfer mit Sauerstoff unter Wärmezufuhr.“			

26 **Abb. 1: Interventions- und Testdesign**

Treatment-Check

Die Intervention sowie die Testungen wurden von der Testleitung und einer geschulten studentischen Hilfskraft manualbasiert und standardisiert durchgeführt. Beide Personen haben nach den Durchführungen Abweichungen vom Manual protokolliert, um die Vergleichbarkeit der Interventionen und Testungen sicherzustellen. Nach Prüfung der Dokumentation kann angenommen werden, dass die Interventionen und Testungen gemäß den Vorgaben durchgeführt wurden.

Mediator- und Kontrollvariablen

Zur Beantwortung der zweiten Forschungsfrage wurde die von den Schülerinnen und Schülern wahrgenommene Unterstützung durch die Rückmeldungen als Mediatorvariable erhoben (adaptiert nach Bürgermeister et al., 2011). Die sechs Aufgaben wurden in 5 Minuten zum Post-Messzeitpunkt eingesetzt ($\alpha = .87$). Darüber hinaus wurden die kognitiven Fähigkeiten ($\alpha_{\text{Version A}} = .87$ und $\alpha_{\text{Version B}} = .85$; Heller & Perleth, 2000) und das Fachwissen zum Thema Metalle ($\alpha = .83$) als Kontrollvariablen erhoben. Für die Erhebung der kognitiven Fähigkeiten wurden 12 Minuten und für das Fachwissen 15 Minuten Testzeit zum Prä-Messzeitpunkt eingeplant. Das Fachwissen wurde zusätzlich auch zum Post- und Follow up-Messzeitpunkt erhoben.

Stichprobe

Die Interventionsstudie wurde mit insgesamt $N = 234$ Schülerinnen und Schülern ($M_{\text{Alter}} = 13,6$; $SD = 0,6$; 49 % weiblich) der 8. Jahrgangsstufe an Gymnasien in Schleswig-Holstein durchgeführt. Von den Schülerinnen und Schülern gehören $n = 73$ Interventionsgruppe 1 ($M_{\text{Alter}} = 13,5$; $SD = 0,6$; 37 % weiblich), $n = 83$ Interventionsgruppe 2 ($M_{\text{Alter}} = 13,4$; $SD = 1,6$; 52 % weiblich) und $n = 78$ Interventionsgruppe 3 ($M_{\text{Alter}} = 13,6$; $SD = 0,6$; 56 % weiblich) an. Die Schülerinnen und Schüler wurden im Klassenverband den Interventionsgruppen zugeordnet. Die Zuordnung der Klassen zu den Interventionsgruppen erfolgte anhand der Ergebnisse zu den kognitiven Fähigkeiten und dem Fachwissen zum Prä-Messzeitpunkt, um gleiche Ausgangsbedingungen in den Interventionsgruppen sicherzustellen. An den teilnehmenden Schulen wurde das Fach Chemie ab der 8. Jahrgangsstufe unterrichtet, sodass die Schülerinnen und Schüler zum Zeitpunkt der Intervention maximal zwölf Wochen lang am Chemieunterricht teilgenommen hatten.

Datenanalyse

Zur Analyse der Lernwirksamkeit der unterschiedlichen Interventionsbedingungen und zur Beantwortung von Forschungsfrage 1 wurden die relativen Lösungshäufigkeiten aus dem NAW-Test sowie aus den Experimentplanungen zum Prä-, Post- und Follow Up-Messzeitpunkt herangezogen. Die Ausgangsbedingungen in den Interventionsgruppen wurden anhand der Daten zum Prä-Messzeitpunkt mittels einfaktorieller Varianzanalysen analysiert. Die Effekte der unterschiedlichen Interventionsbedingungen wurden anschließend anhand von einfaktoriellen Varianzanalysen mit Messwiederholung untersucht. Hierzu wurden die Daten aus den Prä-, Post- und Follow Up-Messzeitpunkten verwendet. Im Rahmen der einfaktoriellen Varianzanalyse mit Messwiederholung wurden der Haupt- und der Interaktionseffekt auf statistische Signifikanz geprüft. Mittels Bonferonni Post hoc-Tests wurden die Effektstärken für die Veränderung innerhalb der Bedingungen über die Messzeitpunkte bestimmt.

Zur Analyse der vermittelnden Funktion der wahrgenommenen Unterstützung der Rückmeldungen auf den Lernzuwachs im Planen von Experimenten unter Anwendung der Variablenkontrollstrategie und damit zur Beantwortung von Forschungsfrage 2 wurden Pfadanalysen und die Bootstrapping-Methode herangezogen (Hayes & Preacher, 2014; Preacher & Hayes, 2008). Beim Bootstrapping wird ein neuer Datensatz durch zufällige Ziehungen von Werten aus dem ursprünglichen Datensatz generiert. Durch solch eine Replikation des ursprünglichen Datensatzes wird eine datenbasierte Verteilung für den statistischen Koeffizienten des indirekten Effekts simuliert. Diese Verteilung muss keine bestimmte Verteilungsform besitzen und kann auch asymmetrisch sein (Efron & Tibshirani, 1986). Die Methode wird daher zur Analyse von Mediatoreffekten und insbesondere indirekten Effekten empfohlen. Außerdem zeigen die Ergebnisse von Simulationsstudien, dass Konfidenzintervalle bessere statistische Eigenschaften besitzen (Efron & Tibshirani, 1986, Hayes & Preacher, 2014, MacKinnon, 2008). Zur Analyse des Mediatoreffekts wurde der hier vorliegende Datensatz 10.000fach repliziert und das Konfidenz-Intervall auf $CI = 98,75$ % festgelegt (vgl. MacKinnon, Lockwood, & Williams, 2004).

Ergebnisse

Zunächst werden die Ergebnisse für Forschungsfrage 1 dargestellt, die die Lernwirksamkeit von unterschiedlichen Rückmeldeformen nach Einführung des Lernziels auf das Planen von Experimenten unter Anwendung der Variablenkontrollstrategie adressiert. Zur Beantwortung von Forschungsfrage 2 wird die vermittelnde Funktion der wahrgenommenen Unterstützung der Rückmeldungen auf den Lernzuwachs beim Planen von Experimenten unter Anwendung der Variablenkontrollstrategie untersucht.

Lernwirksamkeit der Intervention

Zum Prä-Messzeitpunkt liegen in den drei Interventionsgruppen die gleichen Ausgangsbedingungen bezüglich des Wissens zu den naturwissenschaftlichen Arbeitsweisen vor, $F(2, 218) = 1.580, p = .208, \eta^2 = .014$. Vom Prä- zum Post-Messzeitpunkt liegt ein Wissenszuwachs bezüglich dieser Variable vor, $F_{Prä-Post}(1, 180) = 67.354, p < .001, \eta^2 = .272$. Der Zuwachs erfolgt in den drei Interventionsgruppen gleichermaßen, $F_{Prä-Post}(2, 180) = 1.354, p = .261, \eta^2 = .015$. Um den Wissenszuwachs zwischen der Prä- und Post-Testung auf die Intervention zurückführen zu können, werden für die distale abhängige Variable die Daten der drei Interventionsgruppen mit den Daten einer vierten Gruppe verglichen. Die Schülerinnen und Schüler dieser Kontrollgruppe haben nur an den Testungen zu den drei Messzeitpunkten teilgenommen. Zum Prä-Messzeitpunkt weisen die Kontroll- und die Interventionsgruppen ein ähnlich hohes Vorwissen bezüglich der naturwissenschaftlichen Arbeitsweisen auf, $F(3, 274) = 1.027, p = .381, \eta^2 = .011$. Es wird von gleichen Voraussetzungen in allen Gruppen ausgegangen. Für die Kontrollgruppe kann im Unterschied zu den anderen Gruppen kein Wissenszuwachs zwischen Prä- und Post-Messzeitpunkt festgestellt werden, $t(53) = .36, p = .721$. Dementsprechend kann geschlussfolgert werden, dass die Intervention zu einem signifikanten Zuwachs im Wissen über naturwissenschaftliche Arbeitsweisen beigetragen hat.

Neben der distalen Messung der abhängigen Variable erfolgte eine proximale Messung anhand einer Experimentplanung. Zum Prä-Messzeitpunkt weisen die drei Interventionsgruppen die gleichen Ausgangsbedingungen beim Planen von Experimenten auf, $F(2, 216) = 2.554, p = .080, \eta^2 = .023$. Die Schülerinnen und Schüler haben sich anschließend im Rahmen der Intervention im Planen von Experimenten verbessert, $F_{Prä-Post}(1, 190) = 257.346, p < .001, \eta^2 = .575$ (Abb. 2). Zudem haben sich die Schülerinnen und Schüler in Abhängigkeit von ihrer Gruppenbedingung unterschiedlich stark verbessert, $F_{Prä-Post}(2, 190) = 11.989, p < .001, \eta^2 = .112$. Die Effektstärken zeigen, dass die Interventionsgruppe 1 einen höheren Lernzuwachs erfahren hat als die Gruppen 2 und 3 (Abb. 2). Die Gruppe 3 weist den geringsten Lernzuwachs auf. Es kann geschlussfolgert werden, dass Rückmeldungen mit Informationen zum Lernziel, Lernstand und zu nächsten Lernschritten am effektivsten sind.

Abb. 2: Ergebnisse der Experimentplanungen zum Prä-, Post- und Follow Up-Messzeitpunkt in den Interventionsgruppen

Anmerkung. * $p < .05$. ** $p < .01$. *** $p < 0.001$.

Die Experimentplanungen wurden darüber hinaus einer Analyse hinsichtlich unterschiedlicher Qualitätsstufen unterzogen (Abb. 3). Zum Prä-Messzeitpunkt variieren fast alle Schülerinnen und Schüler die unabhängige Variable in ihren Ausprägungen entweder nicht oder lediglich unter unkontrollierten Bedingungen. Wenige Schülerinnen und Schüler variieren die unabhängige Variable unter Angabe einer Messdauer. Zusätzlich zur Messdauer werden keine weiteren Störvariablen zum Prä-Messzeitpunkt berücksichtigt. Demzufolge liegen keine Experimentplanungen mit kontrollierten bzw. stark kontrollierten Bedingungen vor. Zum Post-Messzeitpunkt variieren in der Gruppe 1 und 2 über ein Drittel der Schülerinnen und Schüler die unabhängige Variable unter kontrollierten bzw. stark kontrollierten Bedingungen, in Gruppe 3 nur 14 %. Ein Drittel der Schülerinnen und Schüler variiert die unabhängige Variable jedoch weiterhin nicht. Es lässt sich schlussfolgern, dass die Schülerinnen und Schüler vor der Intervention die unabhängige Variable entweder nicht oder lediglich unter unkontrollierten Bedingungen variieren können. Mehr als ein Drittel der Schülerinnen und Schüler, denen das Lernziel, der Lernstand und nächste Lernschritte bzw. denen das Lernziel und der Lernstand rückgemeldet wurden, können die Variablenkontrollstrategie unter kontrollierten bzw. stark kontrollierten Bedingungen anwenden. Dies gelingt nur wenigen Schülerinnen und Schülern, die ihren Lernstand und die nächsten Lernschritte anhand der aufgezeigten Lernziele selbst eingeschätzt haben.

	Qualität 5	Experiment beinhaltet eine Variation unter stark kontrollierten Bedingungen.
	Qualität 4	Experiment beinhaltet eine Variation unter kontrollierten Bedingungen.
	Qualität 3	Experiment beinhaltet eine Variation nach festgelegter Messdauer.
	Qualität 2	Experiment beinhaltet eine unkontrollierte Variation.
	Qualität 1	Experiment beinhaltet keine Variation der unabhängigen Variable.

Abb. 3: Häufigkeiten in den Qualitätsstufen der Experimentplanung hinsichtlich der Anwendung der Variablenkontrollstrategie innerhalb der Interventionsgruppen

Um die Lernwirksamkeit der Intervention auf deren einzelne Elemente zurückführen zu können, wurde über die beschriebenen Analysen hinaus der Lernzuwachs vom Prä-Messzeitpunkt zur Experimentplanung in der zweiten Interventionsstunde sowie von dieser bis zum Post-Messzeitpunkt untersucht. Zwischen dem Prä-Messzeitpunkt und der ersten Experimentplanung fand die erste Interventionsstunde statt, in der das Lernziel über eine direkte Instruktion eingeführt wurde. Es liegt ein Lernzuwachs vom Prä-Messzeitpunkt bis zur zweiten Interventionsstunde vor, $F_{Prä-2.Int} (1, 194) = 457.201, p < .001, \eta^2 = .702$. Dieser erfolgt in Abhängigkeit von der Gruppenbedingung unterschiedlich, $F_{Prä-2.Int} (2, 194) = 7.363, p = .001, \eta^2 = .071$. Weiterhin liegt eine Veränderung beim Planen von Experimenten von der zweiten Interventionsstunde bis zum Post-Messzeitpunkt

vor, $F_{2.Int-Post}(1, 194) = 3.934, p = .049, \eta^2 = .020$. Diese Veränderung erfolgt in den Interventionsgruppen in unterschiedlichem Maße, $F_{2.Int-Post}(2, 194) = 5.858, p = .003, \eta^2 = .057$. Die Effektstärken für den Zuwachs von der zweiten Interventionsstunde zum Post-Messzeitpunkt zeigen lediglich eine signifikante Veränderung in der Experimentplanung von Interventionsgruppe 3, $\Delta = -.084, SD = .023, p < .001, \eta^2 = .060$. Die Experimentplanung zum Post-Messzeitpunkt weist für diese Gruppe eine geringere Lösungshäufigkeit auf. In den Gruppen 1 und 2 liegt keine signifikante Veränderung in der Experimentplanung vor.

Es kann zusammengefasst werden, dass die Einführung der Merkmale der Experimentplanung einen großen Effekt auf das Planen von Experimenten hat. Dabei erfahren alle drei Interventionsgruppen einen hohen Lernzuwachs durch die Einführung des Lernziels. Die Effektstärken gelten als große Effekte. Die Effektstärke des Haupteffektes von der ersten Experimentplanung zum Post-Messzeitpunkt entspricht einem kleinen Effekt.

Wahrgenommene Unterstützung durch Rückmeldungen

Die Rückmeldungen werden von den Interventionsgruppen signifikant unterschiedlich wahrgenommen, $F(2, 217) = 7.753, p = .001, \eta^2 = .067$. Gruppe 3 nimmt die Rückmeldungen signifikant weniger unterstützend wahr als die Gruppen 1 und 2 (Interventionsgruppe 1 vs. 2: $p = 0.352$; Interventionsgruppe 1 vs. 3: $p = 0.071$; Interventionsgruppe 2 vs. 3: $p < 0.001$).

Vor diesem Hintergrund wird die Rolle der wahrgenommenen Unterstützung mittels Pfadanalyse und Bootstrapping-Methode analysiert. In dem Modell wird der Mediatoreffekt der wahrgenommenen Unterstützung auf den Lernzuwachs beim Planen von Experimenten für die Interventionsgruppen 1 und 2 in Referenz zur Gruppe 3 analysiert. Die wahrgenommene Unterstützung ist in dem verwendeten Modell (Abb. 4) die Mediatorvariable (M). Der Lernzuwachs beim Planen von Experimenten ist die abhängige Variable (Y) und wird anhand der Differenz der Lösungshäufigkeiten in den Experimentplanungen vom Prä- zum Post-Messzeitpunkt operationalisiert. Die Korrelation zwischen den Konstrukten wahrgenommene Unterstützung und Lernzuwachs beim Planen von Experimenten liegt bei $r_s = 0.22, p < .01$. Die Rückmeldung ist die unabhängige Variable. Da für die Rückmeldung drei unterschiedliche Rückmeldeformen vorliegen, muss eine Dummy-Kodierung vorgenommen werden. Hierbei bildet die Interventionsgruppe 3 die Referenzgruppe, da für die Schülerinnen und Schüler dieser Gruppe der geringste Lernzuwachs zu erwarten ist (vgl. Hypothesen; D1: Gruppe 1 = 1, Gruppe 2 = 0; Gruppe 3 = 0; D2: Gruppe 1 = 0, Gruppe 2 = 1, Gruppe 3 = 0).

Im Mediatormodell werden folgende Pfade analysiert: Die Pfade a_1 und a_2 stellen den direkten Effekt auf die wahrgenommene Unterstützung von Gruppe 1 (D₁) und Gruppe 2 (D₂) jeweils in Referenz zur Gruppe 3 dar (Abb. 4). Der Pfad b bildet den direkten Effekt der wahrgenommenen Unterstützung auf den Lernzuwachs in der Experimentplanung ab. Anhand der Pfade c_1 und c_2 wird der direkte Effekt auf den Lernzuwachs in der Experimentplanung von Gruppe 1 (D₁) und Gruppe 2 (D₂) jeweils in Referenz zu Gruppe 3 dargelegt. Die indirekten Effekte auf den Lernzuwachs in der Experimentplanung über die wahrgenommene Unterstützung von Gruppe 1 (D₁) und Gruppe 2 (D₂) jeweils in Referenz zu Gruppe 3 werden anhand der Pfade a_1b sowie a_2b veranschaulicht. Die Güte dieses saturierten Modells wird anhand eines Modellvergleichs mit dem Likelihood-Ratio-Test geschätzt. Hierbei wird das saturierte Modell mit einem reduzierten Modell, das die indirekten Effekte nicht enthält, verglichen (Geiser, 2010). Das saturierte Modell unterscheidet sich signifikant von dem reduzierten Modell ($\chi^2_{diff}(2) = 7.3812, p = 0.025$). Zudem klärt das saturierte Modell eine höhere Varianz als das reduzierte Modell auf ($R^2_{saturiertes\ Modell} = .18; R^2_{reduziertes\ Modell} = .11$). Es kann daher geschlossen werden, dass das saturierte Modell die Daten adäquater abbildet als das reduzierte Modell.

Die Ergebnisse der Pfadanalyse (Abb. 4) zeigen, dass die Rückmeldungen sowohl in Interventionsgruppe 1 ($a_1 = 0.289, p = .016$) als auch in Interventionsgruppe 2 ($a_2 = 0.505, p < .001$) stärker unterstützend wahrgenommen wurden als die Rückmeldungen in Interventionsgruppe 3. Zudem liegt ein signifikanter direkter Effekt der wahrgenommenen Unterstützung auf den Lernzuwachs in den Experimentplanungen vor ($b = 0.949, p = .006$). Das Bootstrapping ergibt, dass der Effekt der Rückmeldung in den Interventionsgruppen 1 und 2 im Vergleich zu der Rückmeldung in Gruppe 3 auf den Lernzuwachs über die wahrgenommene Unterstützung vermittelt wird (D1: $a_1b = 0.274$, BCa-Bootstrap-KI (98,75%) = [0.014, 0.788]; D2: $a_2b = 0.479$, BCa-Bootstrap-KI (98,75%) = [0.100, 0.095]).

- 1
2 a_1 : direkter Effekt lösungsprozessbezogener Rückmeldungen (Gruppe 1) in Referenz zu Selbsteinschätzungen (Gruppe 3; D1)
3 auf die wahrgenommene Unterstützung
4 a_2 : direkter Effekt lernstandbezogener Rückmeldungen (Gruppe 2) in Referenz zu Selbsteinschätzungen (Gruppe 3; D2) auf
5 die wahrgenommene Unterstützung
6 b : direkter Effekt der wahrgenommenen Unterstützung auf den Lernzuwachs in der Experimentplanung
7 c_1 : direkter Effekt lösungsprozessbezogener Rückmeldungen (Gruppe 1) in Referenz zu Selbsteinschätzungen (Gruppe 3; D1)
8 auf den Lernzuwachs in der Experimentplanung
9 c_2 : direkter Effekt lernstandbezogener Rückmeldungen (Gruppe 2) in Referenz zu Selbsteinschätzungen (Gruppe 3; D2) auf
10 den Lernzuwachs in der Experimentplanung

11 Abb. 4: vollständiges Mediatormodell für die wahrgenommene Unterstützung

12 Hinweise: Die berichteten Regressionskoeffizienten sind nicht standardisiert. * $p < .05$. ** $p < .01$. *** $p < 0.001$. Die Pfade der
13 indirekten Effekte a_1b und a_2b werden in der Abbildung nicht dargestellt.

14 Es kann zusammengefasst werden, dass die Schülerinnen und Schüler Rückmeldungen, die das Lernziel, den
15 Lernstand und die nächsten Lernschritte aufzeigen, sowie Rückmeldungen, die über das Lernziel und den
16 Lernstand informieren, stärker unterstützend wahrnehmen als Rückmeldungen, die lediglich das Lernziel
17 enthalten. Zudem wird der Lernzuwachs in den Experimentplanungen über die wahrgenommene Unterstützung
18 der Rückmeldung mediiert.

20 Diskussion der Ergebnisse

21 Anhand der vorliegenden Interventionsstudie wurden unterschiedliche Rückmeldeformen hinsichtlich ihrer
22 Wirksamkeit beim Aufbau der Kompetenz Experimente zu planen untersucht. Im Fokus stand dabei die
23 Anwendung der Variablenkontrollstrategie, welche in bisherigen Untersuchungen als Schwierigkeit von
24 Schülerinnen und Schülern diagnostiziert wurde. Außerdem wurde die wahrgenommene Unterstützung als
25 Mediatorvariable analysiert. Zentraler Befund ist, dass die Rückmeldeformen, die Informationen zum Lernziel,
26 zum Lernstand und zu nächsten Lernschritten liefern, zum höchsten Lernertrag führen. Hierbei ist zu betonen,
27 dass sich die Rückmeldungen im Rahmen der Intervention nur in Kombination mit einer direkten Instruktion zur
28 Einführung des Lernziels als wirksam erwiesen haben. Die wahrgenommene Unterstützung konnte als
29 Mediatorvariable bestätigt werden. Nachfolgend werden die Ergebnisse aus der vorliegenden Studie diskutiert
30 und eingeordnet.

31 Diskussion zu Forschungsfrage 1: Planung von Experimenten unter Anwendung der Variablenkontrollstrategie

32 Als distales Maß für die Kompetenz ein Experiment unter Anwendung der Variablenkontrollstrategie zu planen,
33 wurde das Wissen zu naturwissenschaftlichen Arbeitsweisen erhoben. Die Schülerinnen und Schüler zeigen
34 unabhängig von der Rückmeldeform einen gleich hohen Lernzuwachs in allen drei Interventionsgruppen ($\eta^2 =$
35 $.272 \pm d = 1.222$). Die Schülerinnen und Schüler, die an der Intervention teilgenommen haben, weisen einen
36 signifikant höheren Lernzuwachs im Vergleich zu den Schülerinnen und Schülern aus der Kontrollgruppe auf.

Folglich kann grundsätzlich festgehalten werden, dass die Lernzuwächse auf die Intervention zurückzuführen sind.

Zur Überprüfung der drei Hypothesen müssen die Daten zum proximalen Maß herangezogen werden. Dazu wurden schriftliche Experimentplanungen von den Schülerinnen und Schülern angefertigt und bewertet. Vor Beginn der Intervention können die Schülerinnen und Schüler im Mittel nur ein Viertel der erwarteten Merkmale richtig anwenden. Die Hälfte der Schülerinnen und Schüler nimmt keine Variation der unabhängigen Variable vor und nur wenige Schülerinnen und Schüler variieren die unabhängige Variable unter Angabe einer Messdauer. Keine bzw. keiner der Schülerinnen und Schüler setzt eine Variation der unabhängigen Variable unter kontrollierten bzw. stark kontrollierten Bedingungen in den Experimentplanungen um.

Zum Zeitpunkt der Intervention befanden sich die Schülerinnen und Schüler im Chemieanfangsunterricht. Daher kann angenommen werden, dass die Schülerinnen und Schüler wenig Erfahrung im eigenständigen Planen von Experimenten im Fach Chemie aufweisen. Dementsprechend muss von einer intuitiven Anwendung der Variablenkontrollstrategie ausgegangen werden, bei der die Schülerinnen und Schüler eine Variation der unabhängigen Variable zur Überprüfung der Hypothese umsetzen, jedoch keine weiteren Merkmale der Experimentplanung wie das Konstanthalten von Störvariablen anwenden. Dass ohne Einführung der Variablenkontrollstrategie Schwierigkeiten bei der Variation der unabhängigen Variable unter kontrollierten bzw. stark kontrollierten Bedingungen auftreten, wurde auch in anderen Studien für Schülerinnen und Schüler unterschiedlicher Jahrgangsstufen festgestellt (Arnold et al., 2017; Bullock & Sodian, 2003; Hammann, Phan, Ehmer & Grimm, 2008).

Vom Prä- zum Post-Messzeitpunkt zeigt sich für die Schülerinnen und Schüler ein hoher Lernzuwachs im Planen von Experimenten ($\eta^2 = .58 \triangleq d = 2.326$). Dieser Effekt fällt in Abhängigkeit von der Rückmeldeform unterschiedlich aus. Die Schülerinnen und Schüler, die die lösungsprozessbezogene Rückmeldung erhalten haben, zeigen den höchsten Lernzuwachs ($\eta^2 = .44 \triangleq d = 1.773$), diejenigen, die eine Selbsteinschätzung vornehmen mussten, haben den geringsten Lernzuwachs erzielt ($\eta^2 = .14 \triangleq d = .807$). Es liegt ein mittlerer Effekt vor.

Vor der Planung der Experimente wurde in einer ersten Interventionsstunde das Lernziel eingeführt. Anhand des Unterschieds zwischen der Experimentplanung zum Prä-Messzeitpunkt und der ersten Experimentplanung in der zweiten Interventionsstunde, kann der Effekt der Lernzeleinführung, wie sie bei der formativen Diagnose und Bewertung vorgesehen ist, analysiert werden. Für alle drei Gruppen lässt sich ein großer Effekt für die Einführung des Lernziels festhalten ($.40 \leq \eta^2 \leq .46 \triangleq 1.633 < d < 1.846$). Die Ergebnisse zeigen also, dass die Einführung der Merkmale der Experimentplanung selbst einen hohen Effekt auf das Planen von Experimenten aufweist ($\eta^2 = .702 \triangleq d = 3.070$), während sich für die Rückmeldungen lediglich ein kleiner Lerneffekt auf das Planen von Experimenten ergibt ($\eta^2 = .020 \triangleq d = .286$). Der Effekt der Einführung überlagert scheinbar den Effekt der Rückmeldungen in der Intervention. Die vom Prä- zum Post-Messzeitpunkt analysierte differentielle Wirkung der Rückmeldungen zeigt sich aber auch zwischen der ersten Experimentplanung aus der zweiten Interventionsstunde und dem Post-Messzeitpunkt.

Auch in Meta-Analysen erzielten eine explizite oder implizite Instruktion eines Lernziels deutliche Effekte (explizit: $g = 0.58$, 95% $CI = 0.46 - 0.70$; implizit: $g = 0.65$, 95% $CI = 0.51 - 0.79$, Schwichow et al., 2016). Im Kontext formativer Diagnose und Bewertung scheint dieses Element für die Lernwirksamkeit mitunter entscheidend zu sein (Kluger & DeNisi, 1996). Dahingegen ist die Befundlage zur Wirkung von Rückmeldungen nicht eindeutig (Ross, 1988; Schwichow et al., 2016). Die beiden genannten Meta-Analysen berichten gegensätzliche Befunde. In der hier vorliegenden Studie werden die Ergebnisse von Ross (1988) gestützt, dass Rückmeldungen die Anwendung der Variablenkontrollstrategie fördern können und dass der Effekt von der Rückmeldeform abhängig ist. Im Vergleich zu der lösungsprozessbezogenen Rückmeldung zeigt sich für die lernstandbezogene Rückmeldung ein geringerer Lerneffekt. Es kann geschlussfolgert werden, dass insbesondere die Informationen über nächste Lernschritte lernwirksam sind und dass Hypothese 1a gilt. Es kann vermutet werden, dass der niedrigere Effekt seitens der lernstandbezogenen Rückmeldungen auf die Selbsteinschätzung der nächsten Lernschritte zurückzuführen ist. Die Schülerinnen und Schüler müssen die nächsten Lernschritte anhand eines Vergleichs der Lernziele mit dem rückgemeldeten Lernstand selbst einschätzen. Damit wird die von Hattie und Timperley (2007) abgeleitete Theorie zu den in Rückmeldungen zu gebenden Informationen gestützt, Hypothese 1b aber nicht bestätigt. Der Lernzuwachs fällt bei den Schülerinnen und Schülern der Interventionsgruppe mit Selbsteinschätzung geringer aus als bei den Schülerinnen und Schülern der beiden anderen Gruppen. Der Effekt ist als mittlerer Effekt zu klassifizieren und bestätigt Befunde aus Meta-Analysen ($d = .47$; Falchikov & Boud, 1989). Demnach kann geschlussfolgert werden, dass das Reflektieren des eigenen

Lernstands sowie der nächsten Lernschritte anhand der Lernziele einen positiven Effekt hat, jedoch nicht in solch hohem Maße wie bei den anderen Rückmeldeformen. Hypothese 1c kann damit bestätigt werden.

Trotz der berichteten großen Effekte der Intervention können die Schülerinnen und Schüler zum Post-Messzeitpunkt nur knapp die Hälfte der Merkmale in den Experimentplanungen richtig umsetzen. Eine besondere Schwierigkeit bereitet die Berücksichtigung von Störvariablen sowie die Beschreibung von Kontrollversuchen. Dementsprechend zeigt sich in der Anwendung der Variablenkontrollstrategie, dass ein erheblicher Anteil der Schülerinnen und Schüler in allen drei Interventionsgruppen keine kontrollierten bzw. stark kontrollierten Bedingungen in ihren Experimentplanungen zum Post-Messzeitpunkt umsetzt. Trotz Unterstützungsmaßnahme persistierende Schwierigkeiten berichten ebenfalls Arnold et al. (2017). Die Autoren untersuchten die Lernwirksamkeit von Forschertipps und Concept Cartoons im Hinblick auf das wissenschaftliche Denken mit einer Subskala zum Planen von Experimenten. Dabei stellten sie geringe Lösungswahrscheinlichkeiten bezüglich der Merkmale Messzeiten angeben, Störvariablen berücksichtigen sowie Wiederholungen (bzw. Kontrollversuche) durchführen fest.

Diskussion zu Forschungsfrage 2: wahrgenommene Unterstützung

Forschungsfrage 2 adressiert die vermittelnde Funktion der wahrgenommenen Unterstützung der Rückmeldungen auf das Planen von Experimenten unter Anwendung der Variablenkontrollstrategie. Es wurde die Hypothese 2 aufgestellt, dass die wahrgenommene Unterstützung einen direkten Effekt auf den Lernzuwachs im Planen von Experimenten aufweist und zudem den Effekt der Rückmeldungen auf den Lernzuwachs im Planen von Experimenten mediiert. Es liegt ein signifikanter direkter Effekt von wahrgenommener Unterstützung der Rückmeldungen auf den Lernzuwachs im Planen von Experimenten vor. Zudem ergibt das Bootstrapping einen signifikanten indirekten Effekt von wahrgenommener Unterstützung der Rückmeldungen auf das Planen von Experimenten. Konform mit den Ergebnissen von Wollenschläger et al. (2012) kann geschlossen werden, dass der Effekt der Rückmeldung über die wahrgenommene Unterstützung auf das Planen von Experimenten mediiert wird. Damit kann Hypothese 2 bestätigt werden. Die Analyse der Daten zeigt, dass die Schülerinnen und Schüler in der Intervention die lösungsprozessbezogenen sowie die lernstandbezogenen Rückmeldungen als stärker unterstützend wahrgenommen haben als die Selbsteinschätzungen als Rückmeldungen. Dadurch können lösungsprozessbezogene sowie lernstandbezogene Rückmeldungen das Planen von Experimenten stärker unterstützen als die Selbsteinschätzung. Es kann vermutet werden, dass die Schülerinnen und Schüler bei Selbsteinschätzungen im Vorfeld weitere Übungen benötigen, um ein höheres Verständnis für die Anwendung und Bewertung der Merkmale zu entwickeln (Andrade & Du, 2007; White & Frederiksen, 1998). Ein weiterer Punkt ist, dass die Schülerinnen und Schüler sich möglicherweise als nicht ausreichend kompetent für die Selbsteinschätzung wahrnehmen (Adams & King, 1995). Ferner ist zu hinterfragen, inwiefern die wahrgenommene Unterstützung im Zusammenhang mit einem wahrgenommenen Cognitive Load steht. Bisherige Befunde belegen, dass bei Unterstützungsmaßnahmen ein sogenannter Expertise Reversal Effekt vorliegen kann (Kallweit, 2015, Koenen, 2014). Demnach nehmen Schülerinnen und Schüler mit einem höheren Vorwissen bei Unterstützungsmaßnahmen einen hohen Cognitive Load wahr, der sich mitunter negativ auf ihre Leistungen auswirken kann. Schülerinnen und Schüler mit einem geringen Vorwissen hingegen nehmen eher einen geringen Cognitive Load wahr und weisen einen positiven Leistungsertrag auf. Im Rahmen der vorgestellten Studie wurde der Cognitive Load nicht erhoben, so dass keine evidenzbasierten Aussagen hinsichtlich des Zusammenhangs von wahrgenommener Unterstützung und Cognitive Load getroffen werden können.

Diskussion der Gültigkeit der Untersuchungsbefunde

Bei Interventionsstudien gilt es die ökologische Validität des Interventionsdesigns zu bewerten, um die Generalisierbarkeit der Befunde auf vergleichbare Personengruppen oder Lerngelegenheiten einschätzen zu können (z.B. Bortz & Döring, 2006). Die vorgestellte Untersuchung ist eine quasiexperimentelle Laboruntersuchung. Aufgrund der laborartigen Untersuchungsumgebung (z.B.: Ein Testleiter führt die Studie durch und die Schülerinnen und Schüler arbeiten durchweg in Einzelarbeit.) sowie der Durchführung der Untersuchung in natürlichen Gruppen (Die Zuordnung der Schülerinnen und Schüler zu den Interventionsgruppen erfolgte klassenweise.) sind die externe und interne Validität im Vergleich zu anderen Untersuchungsbedingungen von Interventionsstudien eingeschränkt.

Im Hinblick auf die externe Validität muss vor allem das mit der Intervention adressierte Konstrukt kritisch betrachtet werden. In der Intervention wurde auf die Kompetenz des Planens von Experimenten und den Einfluss von Rückmeldungen auf diese Kompetenz fokussiert. Das Experimentplanen hat sich in empirischen

1 Untersuchungen als eine Schwierigkeit von Schülerinnen und Schülern erwiesen und Rückmeldungen gelten als
2 lernförderlich. Diese Fokussierung bedingt jedoch, dass angrenzende Kompetenzen, wie die Durchführung eines
3 selbstgeplanten Experiments, ausgeblendet werden. Damit lässt sich der beobachtete Effekt zwar eindeutig
4 interpretieren, sprich die interne Validität ist relativ hoch, allerdings lässt sich der Effekt nicht ohne weitere
5 Untersuchungen auf Unterrichtssituationen übertragen. Im Rahmen von regulärem Unterricht muss
6 beispielsweise davon ausgegangen werden, dass Schülerinnen und Schüler mehrere naturwissenschaftliche
7 Denk- und Arbeitsweisen anwenden müssen und dass Schülerinnen und Schüler bei der Durchführung eines
8 Experiments Fehler in der Planung selbst erkennen. Ferner gilt es zu prüfen, inwiefern die berichteten Befunde
9 bei der Wahl anderer inhaltlicher Beispiele aus der Chemie oder aus den anderen naturwissenschaftlichen
10 Domänen Gültigkeit besitzen. Auch hier ist die externe Validität als eingeschränkt zu bewerten.

11 Die laborähnlichen Bedingungen der Untersuchung ermöglichen allerdings die weitgehende Ausschaltung bzw.
12 Kontrolle möglicher Störvariablen, die potenziell auch die abhängige Variable beeinflussen können. Dazu zählt
13 beispielsweise die Arbeit in Kleingruppen beim Experimentieren, in der die soziale Interaktion Lernprozesse
14 beeinflussen kann. Außerdem zählt dazu die Durchführung weiterer Aktivitäten wie das Durchführen von
15 Experimenten, die ebenfalls den Aufbau von Kompetenzen beeinflussen kann. Nicht zuletzt wird die Lehrkraft
16 durch einen Testleiter ersetzt, um etwaige Einflüsse kontrollieren zu können. Die Veränderung der abhängigen
17 Variable kann somit mit hoher Wahrscheinlichkeit ursächlich auf die unabhängige Variable, die
18 unterschiedlichen Rückmeldeformen, zurückgeführt werden. Dadurch wird eine Erklärung der
19 Untersuchungsbefunde ermöglicht und die interne Validität kann als vergleichsweise hoch angesehen werden.
20 Ob die Ergebnisse in dem hier berichteten Maße auch auf andere, „natürliche“ Situationen übertragen werden
21 können, bleibt zu prüfen.

23 **Desiderata**

24 Aus den Ergebnissen der vorliegenden Untersuchung lassen sich drei Desiderata ableiten:

25 Zunächst gilt es in anschließenden Forschungsarbeiten die Unterstützungsmaßnahme selbst weiterzuentwickeln
26 und zu evaluieren. Zwei Merkmale bereiten Schülerinnen und Schülern offensichtlich besondere
27 Schwierigkeiten: Variation der unabhängigen Variable unter Konstanthalten von mehreren Störvariablen und
28 Durchführung von Kontrollversuchen. Beide Merkmale müssen in Unterstützungsmaßnahmen stärker in den
29 Vordergrund gerückt werden. Eine Möglichkeit wären gezielte Trainings zum Umgang mit Rückmeldungen und
30 mit den daraus ableitbaren nächsten Schritten im Lernprozess, die sich genau auf diese beiden Merkmale
31 beziehen. Es muss ferner davon ausgegangen werden, dass die beiden Merkmale durch die Intervention nicht
32 besser gefördert wurden, weil die Intervention zu wenig den Zweck der Merkmale für eine Experimentplanung
33 thematisiert. In Bezug auf deren Wirksamkeit sind folglich Untersuchungen notwendig, die nicht nur die
34 Nutzung der Rückmeldungen optimieren, sondern dabei auch epistemologische Aspekte stärker berücksichtigen
35 (Arnold et al., 2017; Lederman et al., 2014).

36 Da auch nach Durchführung der Intervention ein großer Teil der Schülerinnen und Schüler die
37 Variablenkontrollstrategie nicht unter kontrollierten Bedingungen anwenden kann, ist anzunehmen, dass eine
38 einmalige Einführung der Merkmale zur Experimentplanung für eine adäquate Anwendung der Strategie nicht
39 ausreichend ist. Vielmehr gilt es, die Strategie zu mehreren Zeitpunkten einzuführen und den Aufbau der
40 Kompetenz im Längsschnitt zu verfolgen (Bullock & Sodian, 2003). In diesem Zusammenhang wäre ferner die
41 Frage zu untersuchen, inwieweit eine Adaption einer Einführung an die Voraussetzungen der Schülerinnen und
42 Schüler zielführend ist.

43 Im Rahmen der Intervention wurde in Anlehnung an das geführte forschende Lernen in den drei
44 Interventionsgruppen zu Beginn der zweiten Stunde von der Testleitung eine Fragestellung formuliert und eine
45 zu der Frage passende Hypothese aufgestellt. Die Schülerinnen und Schüler aller drei Interventionsgruppen
46 untersuchten demnach dieselbe Fragestellung sowie dieselbe Hypothese in ihren Experimentplanungen. Eine
47 Erklärung für die relativ ähnlichen Effekte in den Interventionsgruppen kann vermutlich auf das geführte
48 forschende Lernen sowie die stark didaktisch strukturierten Materialien (Forscherbögen, Rückmeldebögen und
49 Reflexionsbögen) zurückgeführt werden. Zu untersuchen wäre demzufolge der Effekt der Rückmeldeformen in
50 unterschiedlich strukturierten Lernumgebungen wie dem offenen forschenden Lernen. Dabei wäre auch zu
51 untersuchen, ob es je nach Fachinhalt unterschiedliche Effekte gibt.

Danksagung

Die Interventionsstudie wurde mit Mitteln des siebten Forschungsrahmenprogramms der Europäischen Kommission (FP7/2007-2013) unter der Fördernummer 321428 finanziert.

Literaturverzeichnis

- Abd-el-Khalick, F., Boujaoude, S., Duschl, R., Lederman, N. G., Mamlok-Naaman, R., Hofstein, A.,... Tuan, H.-L. (2004). Inquiry in science education: International perspectives. *Science Education*, 88(3), 397–419.
- Adams, C. & King, K. (1995). Towards a framework for student self-assessment. *Innovations in Education & Training International*, 32(4), 336–343.
- Andrade, H. & Du, Y. (2007). Student responses to criteria-referenced self-assessment. *Assessment & Evaluation in Higher Education*, 32(2), 159–181.
- Arnold, J., Kremer, K. & Mayer, J. (2017). Scaffolding beim Forschenden Lernen: Eine empirische Untersuchung zur Wirkung von Lernunterstützungen. *Zeitschrift für Didaktik der Naturwissenschaften*, 23(1), 21–37.
- Bangert-Drowns, R. L., Kulik, C.-L. C., Kulik, J. A. & Morgan, M. (1991). The instructional effect of feedback in test-like events. *Review of Educational Research*, 61(2), 213–238.
- Black, P. & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7–75.
- 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.
- Börtz, J. & Döring, N. (2006). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler*. Heidelberg: Springer.
- Bullock, M. & Sodian, B. (2003). Entwicklung des wissenschaftlichen Denkens. In W. Schneider & M. Knopf (Hrsg.), *Entwicklung, Lehren und Lernen* (S. 75–92). Göttingen: Hogrefe.
- Bürgermeister, A., Kampa, M., Rakoczy, K., Harks, B., Besser, M., Klieme, E., ... Leiß, D. (2011). *Dokumentation der Befragungsinstrumente des Laborexperimentes im Projekt 'Conditions and Consequences of Classroom Assessment' (Co²CA)*. Frankfurt, Main: DIPF.
- 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.
- Dolin, J., Black, P., Harlen, W. & Tiberghien, A. (2018). Exploring relations between formative and summative assessment. In J. Dolin & R. Evans (Hrsg.), *Transforming Assessment*. (S. 53-80). Cham, Schweiz: Springer International.
- Efron, B. & Tibshirani, R. (1986). Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statistical Science*, 1(1), 54–77.
- Falchikov, N. & Boud, D. (1989). Student self-assessment in higher education: A meta-analysis. *Review of Educational Research*, 59(4), 395–430.
- Falchikov, N. & Goldfinch, J. (2000). Student peer assessment in higher education: A meta-analysis comparing peer and teacher marks. *Review of Educational Research*, 70(3), 287–322.
- Geiser, C. (2010). *Datenanalyse mit MPlus*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Hadfield, M. (1995). Das Kupfer-Problem. *CHEMKON*, 2(3), 103–106.
- Hammann, M., Phan, T. T. H., Ehmer, M. & Grimm, T. (2008). Assessing pupils' skills in experimentation. *Journal of Biological Education*, 42(2), 66–72.
- Hammann, M., Phan, T. T. H., Ehmer, M. & Bayrhuber, H. (2006). Fehlerfrei Experimentieren. *Der mathematische und naturwissenschaftliche Unterricht*, 59(5), 292–299.
- Hattie, J. & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112.
- Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. London: Routledge.

- 1 Hayes, A. F. & Preacher, K. J. (2014). Statistical mediation analysis with a multicategorical independent
2 variable. *The British Journal of Mathematical and Statistical Psychology*, 67(3), 451–470.
- 3 Heller, K. A. & Perleth, C. (2000). *Kognitiver Fähigkeitstest für 4.-12. Klassen, Revision (KFT 4-12 + R)*.
4 Göttingen: Hogrefe.
- 5 Höttecke, D. (2010). Forschend-entdeckender Physikunterricht: Ein Überblick zu Hintergründen, Chancen und
6 Umsetzungsmöglichkeiten entsprechender Unterrichtskonzeptionen. *Unterricht Physik*, 119, 4–16.
- 7 Höttecke, D. & Rieß, F. (2015). Naturwissenschaftliches Experimentieren im Lichte der jüngeren
8 Wissenschaftsforschung – Auf der Suche nach einem authentischen Experimentbegriff der Fachdidaktik.
9 *Zeitschrift für Didaktik der Naturwissenschaften*, 21, 127–139.
- 10 Kallweit, I. (2015). *Effektivität des Einsatzes von Selbsteinschätzungsbögen im Chemieunterricht der*
11 *Sekundarstufe I – Individuelle Förderung durch selbstreguliertes Lernen*. Berlin: Logos Verlag.
- 12 Kingston, N. & Nash, B. (2011). Formative assessment: A meta-analysis and a call for research. *Educational*
13 *Measurement: Issues and Practice*, 30(4), 28–37.
- 14 Klos, S. (2007). *Kompetenzförderung im naturwissenschaftlichen Anfangsunterricht: Der Einfluss eines*
15 *integrierten Unterrichtskonzepts*. Berlin: Logos.
- 16 Kluger, A. N. & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a
17 meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 119(2), 254–284.
- 18 KMK. (2005). *Beschlüsse der Kultusministerkonferenz: Bildungsstandards im Fach Chemie für den Mittleren*
19 *Schulabschluss*. München: Wolters Kluwer.
- 20 Koenen, J. (2014). *Entwicklung und Evaluation von experimentunterstützten Lösungsbeispielen zur Förderung*
21 *naturwissenschaftlich-experimenteller Arbeitsweisen*. Berlin: Logos Verlag.
- 22 Lederman, J. S., Lederman, N. G., Bartos, S. A., Bartels, S. L., Meyer, A. A. & Schwartz, R. S. (2014).
23 Meaningful assessment of learners' understandings about scientific inquiry. The views about scientific
24 inquiry (VASI) questionnaire. *Journal of Research in Science Teaching*, 51(1), 65–83.
- 25 Lienert, G. & Raatz, U. (1998). *Testaufbau und Testanalyse*. Weinheim: Psychologie Verlags Union.
- 26 MacKinnon, D.P. (2008). *Introduction to Statistical Mediation Analysis*. New York: Taylor & Francis.
- 27 MacKinnon, D. P., Lockwood, C. M. & Williams, J. (2004). Confidence limits for the indirect effect:
28 Distribution of the product and resampling methods. *Multivariate Behavioral Research*, 39(1), 99.
- 29 McMillan, J. H. & Hearn, J. (2008). Student self-assessment: The key to stronger student motivation and higher
30 achievement. *Educational HORIZONS*, 87(1), 40–49.
- 31 Ministerium für Bildung, Wissenschaft, Forschung und Kultur des Landes Schleswig-Holstein (Hrsg.). (1997).
32 *Lehrplan für die Sekundarstufe I der weiterführenden allgemeinbildenden Schulen Hauptschule, Realschule,*
33 *Gymnasium: Chemie*. Verfügbar unter: lehrplan.lernnetz.de/intranet1/index.php?wahl=124
- 34 Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen (2008). *Kernlehrplan für das*
35 *Gymnasium – Sekundarstufe I in Nordrhein-Westfalen. Chemie*. Frechen: Ritterbach.NRC (Hrsg.). (2012). *A*
36 *framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. Washington, D.C.:
37 National Academies Press.
- 38 Preacher, K. J. & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing
39 indirect effects in multiple mediator models. *Behavior Research Methods*, 40(3), 879–891.
- 40 Rönnebeck, S., Bernholt, S. & Ropohl, M. (2016). Searching for a common ground – A literature review of
41 empirical research on scientific inquiry activities. *Studies in Science Education*, 52(2), 161–197.
- 42 Ropohl, M., Scheuermann, H. & Rönnebeck, S. (2015). Diagnostizieren und bewerten mit dem Forscherbogen –
43 Formative Diagnose beim forschenden Lernen. *Naturwissenschaften im Unterricht – Chemie*, 26(149), 40–
44 44.
- 45 Ropohl, M., Rönnebeck, S. & Scheuermann, H. (2015). Naturwissenschaftliche Erkenntnisgewinnung im
46 Chemieunterricht. Das Konzept des forschenden Lernens. *Praxis der Naturwissenschaften – Chemie in der*
47 *Schule*, 64(6), 5–8.
- 48 Ross, J. A. (1988). Controlling variables: A meta-analysis of training studies. *Review of Educational Research*,
49 58(4), 405–437.
- 50 Schulz, A., Wirtz, M. & Starauschek, E. (2012). Das Experiment in den Naturwissenschaften. In W. Rieß, M.
51 Wirtz, B. Barzel & A. Schulz (Hrsg.), *Experimentieren im mathematisch-naturwissenschaftlichen Unterricht.*
52 *Schüler lernen wissenschaftlich denken und arbeiten. (S. 15-38)*. Münster: Waxmann.

- 1 Schwichow, M., Zimmerman, C., Croker, S. & Härtig, H. (2016). What students learn from hands-on activities.
2 *Journal of Research in Science Teaching*, 53(7), 980–1002.
- 3 Siler, S. A. & Klahr, D. (2012). Detecting, classifying and remediating: Children’s explicit and implicit
4 misconceptions about experimental design. In R. W. Proctor & E. J. Capaldi (Hrsg.), *Psychology of science.*
5 *Implicit and explicit processes* (S. 137–180). New York: Oxford University Press.
- 6 Vollmeyer, R. & Rheinberg, F. (2005). A surprising effect of feedback on learning. *Learning and Instruction*,
7 15(6), 589–602.
- 8 Vorholzer, A., Aufschnaiter, C. von & Kirschner, S. (2016). Entwicklung und Erprobung eines Tests zur
9 Erfassung des Verständnisses experimenteller Denk- und Arbeitsweisen. *Zeitschrift für Didaktik der*
10 *Naturwissenschaften*, 22(1), 25–41.
- 11 Wahser, I. & Sumfleth, E. (2008). Training experimenteller Arbeitsweisen zur Unterstützung kooperativer
12 Kleingruppenarbeit im Fach Chemie. *Zeitschrift für Didaktik der Naturwissenschaften*, 14, 219–241.
- 13 Wellnitz, N. & Mayer, J. (2013). Erkenntnismethoden in der Biologie – Entwicklung und Evaluation eines
14 Kompetenzmodells. *Zeitschrift für Didaktik der Naturwissenschaften*, 19, 315–345.
- 15 White, B. Y. & Frederiksen, J. R. (1998). Inquiry, modeling, and metacognition: Making science accessible to all
16 students. *Cognition and Instruction*, 16(1), 3–118.
- 17 Wollenschläger, M., Möller, J. & Harms, U. (2012). Ist kompetenzielles Fremdfeedback überlegen, weil es als
18 effektiver wahrgenommen wird? *Unterrichtswissenschaft*, 40(3), 197–212.
- 19 Zenner, C., Scheuermann, H. & Ropohl, M. (2017). Das Kupfer-Problem 2.0. Untersuchung der
20 Schülervorstellungen zur Oxidation von Kupfer. *Chemie Konkret*, 24(5), 370–380.

DuEPublico

Duisburg-Essen Publications online

UNIVERSITÄT
DUISBURG
ESSEN

Offen im Denken

ub

universitäts
bibliothek

Dieser Text wird via DuEPublico, dem Dokumenten- und Publikationsserver der Universität Duisburg-Essen, zur Verfügung gestellt. Die hier veröffentlichte Version der E-Publikation kann von einer eventuell ebenfalls veröffentlichten Verlagsversion abweichen.

DOI: 10.1007/s40573-018-0080-0

URN: urn:nbn:de:hbz:464-20210531-163622-1

*This is a **post-peer-review, pre-copyedit** version of the article Ropohl, Mathias; Scheuermann, Hilda (2018) Welche Rückmeldungen wirken am besten? Ergebnisse einer empirischen Untersuchung von Rückmeldeformen beim Planen von Experimenten. Published in: Zeitschrift für Didaktik der Naturwissenschaften, 24, S. 151–165, published 24 July 2018. The final authenticated version is available online at: <https://doi.org/10.1007/s40573-018-0080-0>*

Alle Rechte vorbehalten.