

ISABEL WAHNER, ELKE SUMFLETH

Training experimenteller Arbeitsweisen zur Unterstützung kooperativer Kleingruppenarbeit im Fach Chemie

Training of experimental working to support cooperative group work in chemistry education

Zusammenfassung

Die Studie setzt an den Problemen im Bereich des naturwissenschaftlichen Experimentalunterrichts an. Es wird untersucht, ob der Einsatz eines Trainings zum Umgang mit strukturierenden Lernprozesshilfen einen Effekt auf den Lernprozess und den Lernerfolg der SchülerInnen hat. Das Projekt wurde mit insgesamt 172 SchülerInnen der 7. Jahrgangsstufe von drei Gymnasien in Nordrhein-Westfalen als experimentelle Laborstudie durchgeführt. Die Lernwirksamkeit des Trainings und eines zusammenfassenden Lehrervortrags (Feedback) wurden anhand von fachspezifischen Leistungstests und die durch die Hilfen induzierten Veränderungen des Problemlöseprozesses durch eine kategoriengeleitete videobasierte Analyse der Verhaltensbeobachtungen und einen speziell entwickelten Strukturierungstest überprüft. Die Analyse der Test- und Videodaten zeigt einen signifikanten Vorteil der Kombination von Training und Lehrervortrag gegenüber der Kontrollgruppe.

Schlüsselwörter: Feedback, Fehlerkultur, Kleingruppen, naturwissenschaftliche Arbeitsweisen, Strukturierung, Training

Abstract

This study applies to problems in the range of scientific experimental education. It is analysed, whether the use of a training for an adequate dealing with structuring aids has an effect on students' learning process and achievement. Additionally it is tested, to which extend these variables are influenced by a teacher based summary of the results at the end of the experimental phase. The study was conducted with 172 7th grade students as a lab-study with a one-factorial control-group design in three secondary schools in North Rhine-Westphalia. The efficacy of the training and the teacher based summary (feedback) are analysed by a test on chemical knowledge. And the changes in the problem-solving process, related to the given aids, are tested by category based analysis of the video data and a structuring test with respect to the problem-solving process. The analyses of the test and video data show a significant advantage for the combination of training and teacher based lecture compared to the control group.

Keywords: feedback, mistakes, scientific inquiry, small groups, structuring, training

1 Experimentelles Arbeiten

Die Förderung der experimentellen Kompetenz von SchülerInnen wird in den Richtlinien für die Sekundarstufe I gefordert (Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen, 2004) und in den aktuellen Nationalen Bildungsstandards in einem eigenen Kompetenzbereich ‚Erkenntnisgewinnung‘ (Klieme et al., 2003;

KMK, 2005) besonders betont. Weltweit wird dieser Anspruch an den naturwissenschaftlichen Unterricht unter dem Oberbegriff ‚Scientific Literacy‘ diskutiert (Roberts, 2007). Insbesondere aufgrund der Diskussionen der American Association for the Advancement of Science (AAAS) (1993, 2001) und des National Research Council (NRC) (2000) erlangt der Forschungsbereich um ‚Nature of Science‘ (vgl.: Flick & Lederman, 2004;

McComas, 1998) international neue Aktualität¹. Es existieren verschiedene Definitionen von ‚Nature of Science‘ (Lederman, 2007; Abd-El-Khalick & Lederman, 2000; Osborne et al., 2003). Schwartz et al. (2004) weisen ‚Nature of Science‘ acht charakterisierende Teilaspekte zu, während in den Standards der American AAAS (1990, 1993, 2001) beim Theoriekonstrukt ‚Nature of Science‘ drei Hauptaspekte unterschieden werden: ‚Scientific World View‘, ‚Scientific Inquiry‘ und ‚Scientific Enterprise‘. Beiden gemeinsam ist die Forderung nach einem reflektierten Umgang mit Experimenten. In dieser Studie steht ‚Scientific Inquiry‘ im Mittelpunkt. ‚Scientific Inquiry‘ wird im deutschsprachigen Raum mit naturwissenschaftlichen Arbeitsweisen gleichgesetzt. Ähnlich wie bei der Definition von ‚Nature of Science‘ liegt auch bei ‚Scientific Inquiry‘ eine Vielzahl von Definitionen vor (vgl.: Martin-Hauser, 2002; Minstrell & van Zee, 2000; Barrow, 2006). Die Ausführungen der AAAS (2001) und des NRC (2000) sind die momentan bekanntesten und aktuellsten, wonach es sich bei ‚Inquiry‘ sowohl um ein Lernziel als auch um einen Lehransatz handelt (Brown et al., 2006). Innerhalb dieses Projekts wird ‚Inquiry‘ als Lernziel verstanden und verwendet. Auch die Inhaltsbereiche von ‚Scientific Inquiry‘ lassen sich weiter unterteilen (vgl.: AAAS, 2001):

- ‚Evidence and Reasoning‘: Experimentalbeobachtungen, Bedeutsamkeit von qualitativen Beobachtungen, Wechselwirkung von Beurteilung und logischer Begründung (...)
- ‚Scientific Investigations‘: verschiedene Untersuchungstypen zur Überprüfung von Hypothesen, isolierende Variablenkontrolle (...)
- ‚Scientific Theories‘: Verständnis von Theorien, deren Anwendung und Veränderung (...)

Die experimentellen Arbeitsweisen gehören somit zu dem Bereich der ‚Scientific Investigations‘. In Anlehnung an den Problemlöseprozess nach Klahr (2000) wird der idealisierte experimentelle Problemlöseprozess in dieser Arbeit wie folgt strukturiert: Ausgehend von einer Problemstellung sind Hypothesen zur Problemlösung zu formulieren. Diese sind anhand eines sinnvollen Experiments zu prüfen. Die experimentellen Beobachtungen müssen im Anschluss im Sinne einer Schlussfolgerung auf die Hypothese bezogen werden und diese verifizieren oder falsifizieren. Werden diese Erkenntnisse dann im Folgenden in das bestehende Vorwissen integriert, wird neues Wissen aufgebaut (vgl.: Niedderer, 1996; Newell & Simon, 1972; Dörner, 1976; Wirth & Leutner, 2006). Da die Ergebnisse eines Vorgängerprojekts gezeigt haben, dass SchülerInnen der 7. Jahrgangsstufe keine theoriegeleiteten Überlegungen zur Hypothesenbildung und -überprüfung anstellen, wird innerhalb dieses Projekts von Ideen gesprochen (Walpuski, 2006).

Experimentelles Arbeiten im Sinne des naturwissenschaftlichen Problemlösens muss explizit unterrichtet werden, es wird nicht durch Nachahmung bzw. Abarbeiten einer Vorlage erlangt (Bell et al., 2003). Sinn und Zweck der Vorgehensweise müssen im Unterricht thematisiert werden (Schwartz et al., 2004). Lederman & Lederman (2006) gehen über diesen Aspekt hinaus, wenn sie die theoretische Vermittlung durch die praktische Umsetzung ergänzen und dieser eine entscheidende Rolle im Verstehensprozess zuweisen (Gallagher-Bolos & Smithenry, 2004; Lederman, 2007). Dabei ist es unumgänglich, dass die Lehrenden selbst über entsprechendes Wissen über ‚Nature of Science‘ und ‚Scientific Inquiry‘ verfügen. Auch hier bestehen große Defizite (Lederman & Lederman, 2006; Brown et

¹ Eine historische Abhandlung findet sich bei Atkin & Black (2007), Barrow (2006) oder Bybee (1997)

al., 2006; Akerson & Volrich, 2006; Akerson & Hanuscin, 2007).

In aktuellen Schüler- und Lehrermaterialien für die Unterstufe bzw. für den Anfangsunterricht in den Naturwissenschaften findet sich der beschriebene Problemlöseprozess in verschiedenen Komplexitätsabstufungen (vgl.: Hübinger & Sumfleth, 2006; Stäudel et al., 2002; Bäurle et al., 2004). Allen sind die Hauptschritte ‚Idee‘, ‚Experiment‘ und ‚Schlussfolgerung‘ gemein. Bisher existieren jedoch kaum empirische Untersuchungen über die Vermittlung dieser im Schulunterricht. Es können jedoch Hinweise darauf gegeben werden, dass die Vermittlung des Prozesses zu einem Lernzuwachs bei den SchülerInnen führt (Bell et al., 2003; Schwartz et al., 2004).

2 Fehlerkultur

Schon Weimer (1925) charakterisiert Fehler als in der Schule unerwünschte Störungen, so dass ein Lernen aus Fehlern nur schwer möglich ist (Glück, 1999). Erst in den letzten Jahren wurde die Thematik der Fehlerkultur – insbesondere durch die Arbeitsgruppe um Oser (vgl.: z.B. Oser & Spychiger, 2005) und im SINUS-Projekt (Modul 3; Müller & Nieswandt, 1999) – aufgegriffen. Diese Arbeiten betonen, dass Fehler insbesondere beim entdeckenden Lernen (Hammerer, 2001; Kahl, 2002) zum Lernen dazu gehören (Mathan & Koedinger, 2005). Sie haben sowohl eine präventive (z.B. Vermeidung des Wiederholens von Fehlern) als auch eine produktive Wirkung (Lernzuwachs). So gilt ein Fehler als ein von einer Norm abweichender Sachverhalt oder auch Prozess, durch welchen es überhaupt erst möglich wird, einen richtigen normbezogenen Sachverhalt oder Prozess in seinen Abgrenzungen zu erkennen (Rollett, 1999).

Fehler werden i.d.R. als Makel oder Defizit empfunden, welche zu vermeiden sind. SchülerInnen erkennen auch deshalb bei Fehlern keine Lernchancen. Es erscheint ihnen unlogisch, dass aus einem „negativen“ Fehler ein

„positiver“ Wissenszuwachs resultieren soll. Fehler können aber sehr wohl positiv und stimulierend erlebt werden. Spychiger et al. (1999) greifen diesen Aspekt in ihrer Theorie des negativen Wissens (s.u.) auf.

Einen zentralen Punkt in den Arbeiten von Oser (Oser & Spychiger, 2005; Oser et al., 1997b, 1998, 1999b und Spychiger et al., 2000) bildet die Theorie des negativen Wissens. ‚Negativ‘ ist nicht im moralisch-wertenden Sinne zu verstehen oder als ‚zu Vermeidendes‘, sondern vielmehr als Gegenstück zum ‚positiven‘ Wissen, welches richtige Abläufe und Resultate beinhaltet (Oser et al., 1999a). Oser & Spychiger (2000) gehen davon aus, dass negatives Wissen durch das Machen von Fehlern gebildet wird. Das negative Wissen ist demnach „das Wissen um das, wie etwas nicht ist (deklarativ) oder nicht funktioniert (prozedural)“ (Oser et al.; 1999a, 17). Folglich ist das negative Wissen elementar bei Prozessen des Problemlösens, der Routinebildung oder des Konzeptaufbaus. Negatives Wissen kann allerdings nicht nur durch eigenes Fehlermachen in realen oder simulierten Situationen erworben werden, sondern auch durch das Erkennen von fremden Fehlersituationen sowie durch den Transfer von positivem Wissen (Oser et al., 1999a). Bei einer gelungenen Wissensanwendung auf eine ähnliche Problemsituation wissen die SchülerInnen auch um das jeweils Falsche.

Oser et al. (1999a) beschreiben mit dem Begriff ‚Fehlerkultur‘ sinnvolles Ausprobieren und das Abarbeiten falscher Ansätze. Damit ist nicht ein Ausprobieren im Sinne von ‚zufälligem Zusammenschütten‘ ohne Variablenkontrolle gemeint, sondern vielmehr ein hypothesengeleitetes Ausprobieren gemäß dem weiter oben beschriebenen Problemlöseprozess und der Verifizierung/Falsifizierung von Hypothesen. Die SchülerInnen sollen für (eigene) Fehler sensibler werden, diese analysieren, korrigieren, den Sachverhalt/Prozess richtig wiederholen und letztendlich lernen den jeweiligen Fehler zu vermeiden. Der richtige Umgang mit Fehlern soll demnach den Lernprozess und

die Unterrichtsqualität optimieren (Oser et al., 1997a; Spychiger et al., 1997).

3 Feedback

Feedback kann SchülerInnen helfen, Fehler zu erkennen und zu analysieren, Ursachen hierfür zu suchen und schließlich negatives Wissen aufzubauen. Es ist dabei unumgänglich, dass die Lernenden wissen, dass Fehler akzeptiert werden – zum Teil sogar erwünscht sind – und mit Hilfe von Feedback eine neue Erkenntnis erlangt werden kann. Dabei müssen bezüglich der Bewertung von Fehlern Unterschiede zwischen Lehr- und Prüfungssituationen verdeutlicht werden (Jacobs, 2002; Hammerer, 2001).

Nach Hattie & Timpeley (2007) ist jedes Feedback mit der Absicht verbunden, die Differenz zwischen dem situational bestehenden Wissen und dem erwünschten Lernziel zu verringern. Hierbei kann Feedback drei grundlegende Funktionen haben: Die Elimination der falschen Antwort sowie die Substitution und Stärkung der korrekten Antwort (Kulhavy et al., 1990). Die Wirkung des Feedbacks hängt dabei insbesondere von der Bereitschaft der SchülerInnen ab, die Informationsangebote zu nutzen. Unter Einbezug von Jacobs (2002) und Narciss (2004) können verschiedene Feedbackarten unterschieden werden (Bangert-Drowns et al., 1991; Narciss & Huth, 2004; Butler & Winne, 1995; Pommer, 2004):

a) Knowledge of result (KOR)

Den SchülerInnen wird mitgeteilt, ob ihre Antwort richtig oder falsch ist. Im Falle einer falschen Antwort bleibt die richtige Antwort unbekannt.

b) Knowledge of correct result (KCR) / korrektives Feedback

Den SchülerInnen wird mitgeteilt, ob ihre Antwort richtig oder falsch ist, und sie erhalten zusätzlich die richtige Antwort.

c) Answer until correct (AUC)

KOR wird mehrmals wiederholt, bis die SchülerInnen die korrekte Antwort selbst finden.

d) Elaborated feedback

Den SchülerInnen wird mitgeteilt, warum ihre Antwort richtig bzw. falsch ist. Hierbei lassen sich verschiedene Arten unterscheiden:

1. task specific elaboration

Den SchülerInnen wird die richtige Antwort mitgeteilt, jedoch nicht erläutert. Der Begriff Elaboration wird hier fälschlich angewandt.

2. instruction based elaboration

Die Erläuterung der Antwort basiert auf dem ursprünglichen Lernmaterial.

3. extra-instructional elaboration

In der Erläuterung werden den SchülerInnen neue Informationen zur Verfügung gestellt.

Beim elaborierten Feedback ist es unumgänglich, Informationen zu geben, welche die korrekte Lösung erläutern. Laut Rollett (1999) ist es dabei von großer Wichtigkeit, dass der Lehrer darauf hinweist, „welche Aspekte des Lösungsverhaltens zum Erfolg führten“ (ebd., 79).

Das ‚informative tutorielle Feedback‘ (Narciss, 2004; Narciss & Huth, 2004) kann als weitere Unterform des elaborierten Feedbacks angesehen werden. Die Lehrperson dient hier lediglich der Absicherung des Schülers bzw. der Schülerin und der Unterstützung des Lernprozesses. Die gegebenen Hilfestellungen können variabel in Umfang und Komplexität eingesetzt werden.

Eine Metaanalyse von Bangert-Drowns et al. (1991) belegt bessere Lernleistungen der SchülerInnen durch die Anwendung von Feedback. Auch andere Studien zeigen hohe Effektstärken des korrektiven Feedbacks (Lysakowsky & Walberg, 1982; Jacobs, 1998). Allerdings wird in einigen Studien auch ein signifikanter Nachteil von Feedback gegenüber No-Feedback festgestellt (z.B. Pridemore & Klein, 1995).

Hinsichtlich des Erfolgs elaborierten Feedbacks gegenüber dem korrektiven Feedback liegen inkonsistente Ergebnisse vor (Hattie & Timperley, 2007). So zeigt die Studie von Merrill (1987) keine Verbesserung der Lernerleistung durch Nutzung des elaborierten Feedbacks im Vergleich zur Nutzung von korrektivem Feedback (Jacobs, 2002), während bei Anderson et al. (1995) durch den Einsatz von elaboriertem Feedback weniger Fehler auftreten und die auftretenden Fehler schneller korrigiert werden. Auch Farquhar (1995) kann nachweisen, dass elaboriertes Feedback korrektivem Feedback hinsichtlich Lernzeit und Lernerfolg überlegen ist, allerdings nur dann, wenn das Feedback unmittelbar und nicht verzögert gegeben wird. Die Frage nach dem Erfolg von elaboriertem gegenüber korrektivem Feedback kann bisher folglich nicht endgültig beantwortet werden.

4 Ziele der Untersuchung

Die Ergebnisse der o.g. vorangegangenen Studie zeigen, dass Feedbackmaßnahmen bei der experimentellen Kleingruppenarbeit einen nachweisbar positiven Effekt auf die Schülerleistung haben, Strukturierungsmaßnahmen jedoch weitgehend wirkungslos blieben (Walpuski & Sumfleth, 2007; Walpuski, 2006). Durch die Entwicklung eines Trainings zum Umgang mit den vorgegebenen prozessunterstützenden Strukturierungshilfen wird in diesem Projekt versucht, den Lernerfolg der SchülerInnen durch einen adäquaten Umgang mit den Strukturierungshilfen aber auch mit den eigenen Fehlern zu optimieren. Die Lernwirksamkeit der Maßnahmen wird durch eine kategoriengeleitete Analyse der Inhaltshandlungen der SchülerInnen auf der Basis von Videodaten

und durch Leistungskontrolle anhand von paper-and-pencil-Verfahren erfasst. Es geht somit um folgende Forschungsfragen:

- Werden durch den Einsatz des Strukturierungstrainings Quantität und Art der Schülerfehler geändert?
- Kann durch das Training ein strukturiertes experimentelles Arbeiten der SchülerInnen gefördert werden?
- Lässt sich durch die Einführung eines Strukturierungstrainings der Lernerfolg der Kleingruppen durch den Umgang mit den angebotenen strukturierenden Lernprozesshilfen steigern?

5 Intervention und Design

An fünf aufeinander folgenden Schultagen hatten die SchülerInnen fünf aufeinander aufbauende Aufgaben aus dem Themenbereich Säure-Base im Anschluss an ihren regulären Unterricht zu bearbeiten. Das Themengebiet beginnt dabei mit der Einteilung von Lösungen in sauer, basisch und neutral und endet mit der Neutralisation von Saurer Regen. Um die Aufgaben bearbeiten und lösen zu können, erhielten die SchülerInnen Interaktionsboxen (Rumann, 2005), die Geräte, Chemikalien und inhaltliche Hilfen sowie die Aufgabenstellungen beinhalteten. Außerdem hatten alle SchülerInnen die Möglichkeit, Feedback vom Lehrer einzufordern: ob ihr Experiment geeignet war, ihre Idee zu überprüfen und ob sie das Experiment richtig durchgeführt hatten.

Zusätzlich erhielten alle SchülerInnen eine Strukturierungshilfe, die den naturwissenschaftlichen Problemlöseprozess abbildet². Die bereits in der Vorgängerstudie (Walpuski & Sumfleth, 2007) eingesetzte Strukturierungshilfe wurde optimiert (Abb. 1). Die ein-

²Es ist bekannt, dass dieser Prozess nicht der einzig mögliche und richtige Weg ist (vgl.: Abell & Lederman, 2007). Um die kognitiven Fähigkeiten der SchülerInnen jedoch nicht zu überfordern und um einen ersten Schritt in Richtung von ausgeprägtem metakognitivem Wissen über naturwissenschaftliche Arbeitsweisen zu machen, wurde den SchülerInnen in dieser Studie lediglich dieser eine Prozessablauf vermittelt.

zelen Schritte sind nun überwiegend linear angeordnet, um den Verstehensprozess der SchülerInnen zu erleichtern; außerdem wurden kindgerechte Abbildungen zur Symbolisierung der einzelnen Schritte hinzugefügt. Es wurde zunächst eine Pilotstudie durchgeführt, um die verschiedenen neu entwi-

ckelten Testinstrumente (z.B. Fachtest) sowie die beiden Interventionen Strukturierungstraining und den Lehrervortrag (s.u.) in einem 2*2-Design zu evaluieren. Die Schüler mit der Intervention Strukturierungstraining erhielten zu Beginn der ersten drei Kleingruppenarbeitsphasen ein Training zum

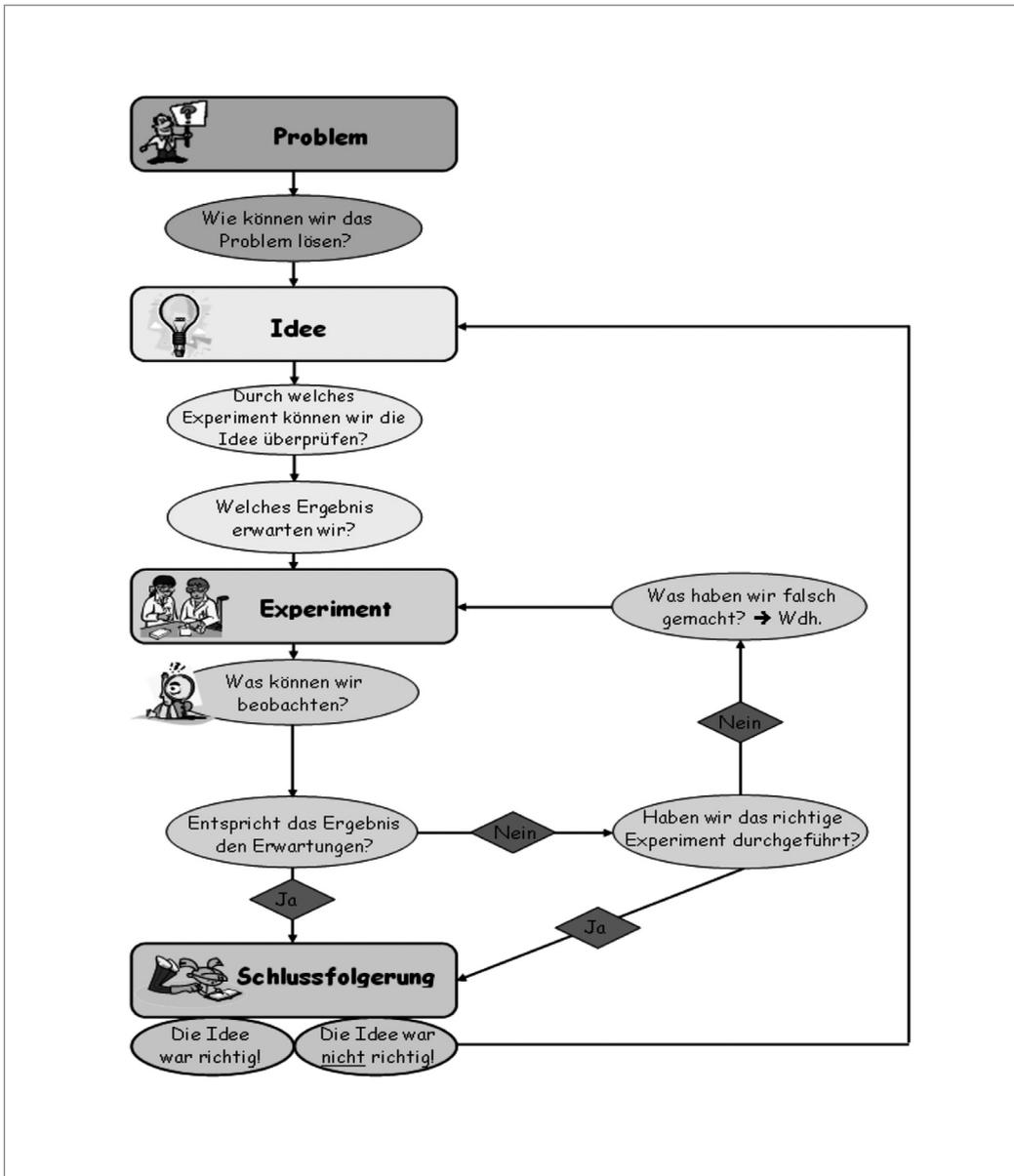


Abb. 1: Strukturierungshilfe

Umgang mit der Strukturierungshilfe, welche den naturwissenschaftlichen Problemlöseprozess abbildet. Die Intervention Lehrervortrag wurde zusätzlich erhoben, um zu analysieren, inwieweit eine Ergebnissicherung durch den Lehrer am Ende der Kleingruppenarbeitsphasen die Schülerleistung beeinflusst und ggf. die Beeinflussung der Lernleistung durch einen adäquaten Umgang mit der Strukturierungshilfe unterstützt. Mit Hilfe einer standardisierten Power-Point-Präsentation wurde den SchülerInnen der Inhalt der zuvor durchgeführten Kleingruppenarbeitsphase in Form einer Zusammenfassung von ca. 5 Minuten Dauer vorgelesen. Diese Form des Feedbacks entspricht dabei dem elaborierten Feedback (instruction based elaboration, vgl.: Kapitel 3).

Die Ergebnisse der Pilotstudie zeigen im Fachtest, dass beide Interventionen einen positiven Effekt auf die Lernleistung der SchülerInnen haben, insbesondere in Kombination. Allerdings unterscheiden sich die Treatmentgruppen, in denen nur die Intervention Strukturierungstraining oder nur die Intervention Lehrervortrag durchgeführt werden, in den Fachtestergebnissen nicht voneinander. Hier wurden aufgrund dieser Ergebnisse auch in der Hauptstudie keine Unterschiede erwartet. Deshalb wurden in der Hauptstudie nur drei Treatments durchgeführt. So ergibt sich ein 2*2-1 Untersuchungsdesign (Abb. 2).

Treatment A fungiert in der Hauptstudie als Kontrollgruppe. Die SchülerInnen erhalten

lediglich die o.g. Interaktionsbox und die Strukturierungshilfe. Die SchülerInnen des Treatments B erhalten zusätzlich im Anschluss an die jeweilige Kleingruppenarbeitsphase den standardisierten Lehrervortrag. In Treatment C werden sowohl die Intervention Lehrervortrag als auch die Intervention Strukturierungstraining durchgeführt. Die Intervention Strukturierungstraining beinhaltet dabei ein Training zum Umgang mit der Strukturierungshilfe, das zu Beginn der ersten drei Kleingruppenarbeitsphasen durchgeführt wird, wobei jede Trainingseinheit etwa 20 Minuten dauert. Den SchülerInnen wird mit Hilfe von Beispielen und verschiedenen Aufgaben der abgebildete naturwissenschaftliche Problemlöseprozess erläutert und verdeutlicht.

In der ersten Trainingseinheit wird mit den SchülerInnen die Strukturierungshilfe Schritt für Schritt konstruiert (vgl. Abb. 3). Die Konstruktion erfolgt anhand eines einfachen Beispiels aus der Physik: Die Entdeckung der Glühbirne durch Edison. Hierbei handelt es sich lediglich um grundlegende Wissensstrukturen. Die SchülerInnen sollen nicht zusätzlich weitere Lernleistung erbringen müssen. Aus demselben Grund wurde auch kein Beispiel aus der Chemie gewählt.

In der zweiten Trainingseinheit werden insbesondere die notwendige Passung von Idee und Experiment thematisiert. Dies bedeutet, dass ein Experiment geeignet sein muss, um eine Idee zu überprüfen. Die SchülerInnen sollen für die Beachtung dieses Zusammen-

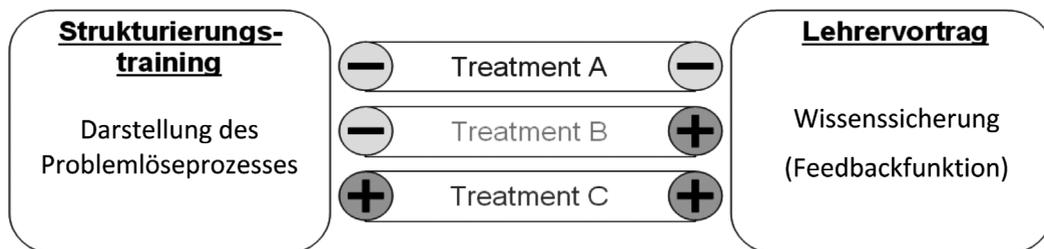


Abb. 2: Treatments der Hauptstudie

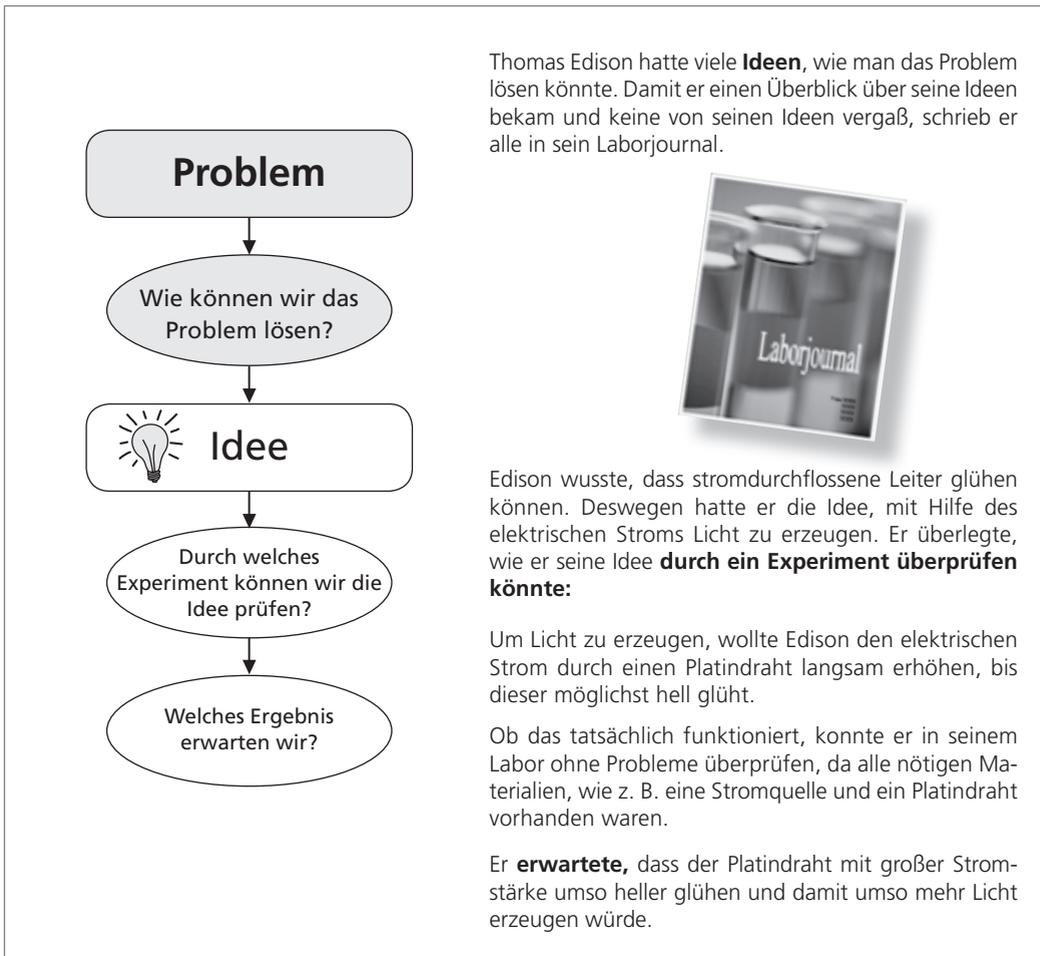


Abb. 3: Beispiel aus der ersten Trainingseinheit

hangs während ihrer eigenen Experimentierphase sensibilisiert werden. In der letzten Trainingseinheit wird der Unterschied von „Beobachtung“ und „Schlussfolgerung“ besprochen.

In allen Treatments haben die SchülerInnen innerhalb von 20 Minuten die jeweils gestellte Aufgabe experimentell zu bearbeiten. Weitere Instruktionen (z.B.: Hinweis auf die Führung eines Laborjournals) nehmen zusätzlich nicht mehr als 5 Minuten ein.

Das Strukturierungstraining und der Lehrervortrag wurden mit Hilfe von standardisierten Manualen durch den Untersuchungsleiter durchgeführt. Die Kleingruppen wurden neben dem Untersuchungsleiter auch von entsprechend instruierten

studentischen Hilfskräften angeleitet. Auch hier wurde ein entsprechendes standardisiertes Manual eingesetzt. Es ist eine sehr hohe Objektivität gegeben. Gleiches gilt für die Hauptstudie.

Es wird erwartet, dass die SchülerInnen der Treatmentgruppe C besser abschneiden als diejenigen der beiden anderen Treatments. Außerdem wird davon ausgegangen, dass die SchülerInnen des Treatments A (Kontrollgruppe) schlechter abschneiden als die übrigen.

Die Hauptstudie wurde an drei Gymnasien in Nordrhein-Westfalen mit insgesamt 172 SchülerInnen der 7. Jahrgangsstufe als experimentelle Laborstudie durchgeführt (Abb. 4).

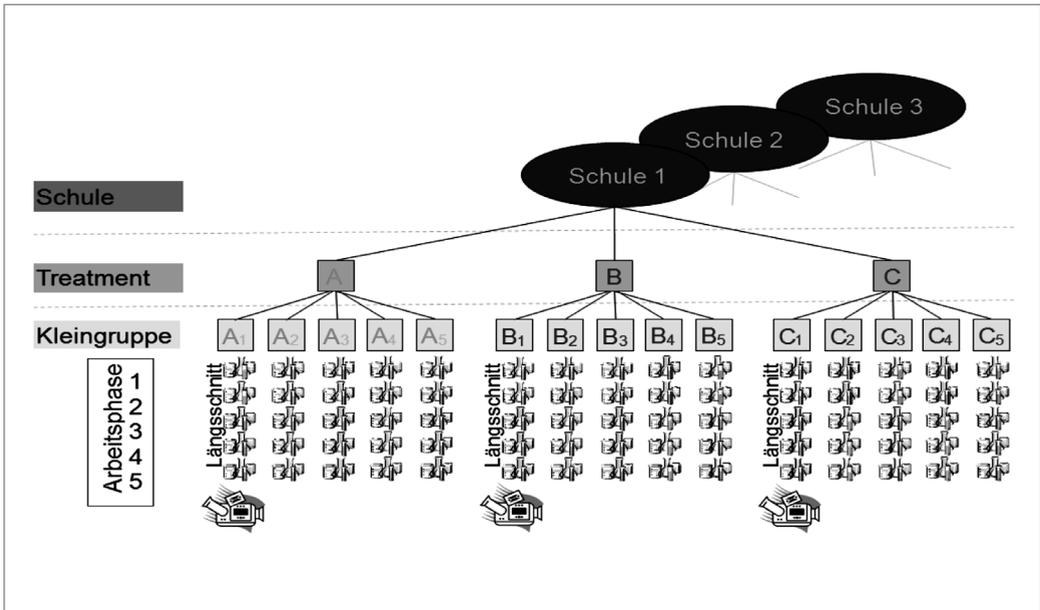


Abb. 4: Durchführung der Hauptstudie

Alle SchülerInnen nahmen freiwillig an der Studie teil und wurden durch einen geringen finanziellen Betrag für den zeitlichen Aufwand von ca. 1-2 Stunden pro Tag entschädigt. Die Treatments wurden an jeweils einer Schule parallel durchgeführt, wobei jeder Gruppe ein eigener Schulraum vorbehalten war. Mittels F-Test (Within-Faktor) werden die statistische Power und die nachweisbaren Effekte berechnet. Auf Ebene der Probanden ($N = 172$) können kleine bis mittlere Effekte ($f^2 = .06$) mit einer statistischen

Power von $1-\beta = .83$ ermittelt werden (Erdfelder et al., 1996).

Die Untersuchung wurde mit einem Pre-Post-Follow-up-Test-Design durchgeführt. Eine Woche vor Untersuchungsbeginn wurde die Pre-Testung durchgeführt. Die Post-Testung erfolgte eine Woche und die Follow-up-Testung ca. drei Monate nach der Projektdurchführung. Tabelle 1 gibt einen Überblick darüber, zu welchen Zeitpunkten die Testinstrumente eingesetzt werden, über deren Ergebnisse hier berichtet wird:

Tab. 1: Testübersicht, Hauptstudie

Test	Pre	Begleitend	Post	Follow-up
KFT	X			
Sozialfragebogen	X			
Fachtest (MC)	X		X	X
Strukturierungstest	X		X	X
Videographierung		X		

Um die Treatmentgruppen vergleichbar zu halten, wurden sie mit Hilfe eines kognitiven Fähigkeitentest (KFT) und den Ergebnissen des Fachtests zum Pre-Zeitpunkt balanciert. Beim Fachtest handelt es sich um einen anhand der Inhalte der Kleingruppenarbeitsphasen neu entwickelten Leistungstest mit insgesamt 21 Items, wobei jedes Item vier Antwortmöglichkeiten aufweist. Die Anzahl der Distraktoren und Attraktoren variiert zwischen den einzelnen Items.

Mit Bezug zum Strukturierungstraining wurde ein Strukturierungstest entwickelt, mit dem der Wissensstand der SchülerInnen über Strategien zur Anwendung der vorgegebenen Strukturierungshilfen geprüft wird. Aufgrund fehlender Literatur bzw. Forschungsergebnisse und der o.g. Begrenzung des dargestellten Problemlöseprozesses musste der Strukturierungstest neu entwickelt werden. Der Test gliedert sich in zwei Teile. Im ersten Teil des Tests wird zunächst das Wissen über Zusammenhänge zwischen den Schritten des Problemlöseprozesses erhoben, wie z.B. der Zusammenhang von Idee und Experiment. Hierzu wurden 36 Aussagen formuliert, die zu bejahen oder zu verneinen sind. So müssen sich die SchülerInnen z.B. entscheiden, ob sie der Aussage „Zeigt ein Experiment ein unerwartetes Ergebnis, obwohl das Experiment richtig durchgeführt wurde, dann war die Idee falsch“ zustimmen können oder nicht. Im zweiten Teil des Strukturierungstests müssen die SchülerInnen 20 verschiedene Aussagen den Schritten des Problemlöseprozesses zuordnen. Die SchülerInnen haben dabei die Wahl zwischen ‚Beobachtung‘, ‚Experiment‘, ‚Idee‘, ‚Problem‘ und ‚Schlussfolgerung‘.

Weitere prozessbegleitende Daten liefern die Videoaufzeichnungen der Arbeitsphasen einzelner Kleingruppen (s. Abb. 2), die kategoriengeleitet ausgewertet wurden und so Hinweise auf die Umsetzung der Strukturierungsmaßnahmen liefern. Hierbei wurde der Schwerpunkt auf die Betrachtung auftretender Fehler gelegt, die mit Hilfe eines entsprechenden Manuals kodiert und ausgewertet wurden.

6 Ergebnisse

6.1 Multiple-Choice-Fachtest

Der Fachtest mit 21 Items weist mit einem Alpha-Koeffizienten von .881 zum Pre- und von .863 zum Post-Zeitpunkt eine ausreichende Reliabilität auf (vgl.: Bortz 2004). Der Vergleich von Pre- und Posttestdaten zeigt, dass in allen Gruppen ein hochsignifikanter Lernzuwachs vorliegt ($p < .001$). Die Lernwirksamkeit der Lernumgebung ist somit analog zu Rumann (2005) und Walpuski (2006) nachgewiesen.

Für die weitere Analyse der Testdaten werden die Residualwerte der einzelnen Treatmentgruppen herangezogen. Mittels ANOVA wird nachgewiesen, dass sich die Mittelwerte der Residuen zwischen den Treatmentgruppen hochsignifikant von einander unterscheiden ($F(2;165) = 17,063$, $p < .001$, $\eta^2 = .174$).

Die Kontrollgruppe A und die Treatmentgruppe B bleiben hinter der erwarteten Lernleistung zurück, während die Treatmentgruppe C die Erwartungen übertrifft. Die residualen Mittelwerte der Treatmentgruppe C unterscheiden sich hochsignifikant von Treatmentgruppe A und signifikant von B. Auch die Mittelwerte der Treatmentgruppen A und B unterscheiden sich signifikant (Abb. 5). Folglich ist die Intervention mit zusätzlichem Lehrervortrag und Strukturierungstraining lernwirksamer als die Intervention, die nur den ergänzenden Lehrervortrag enthält, und beide sind lernwirksamer als die Basislernumgebung.

6.2 Strukturierung

Der erste Teil des Strukturierungstests weist ausreichende Testgütekriterien für einen Gruppenvergleich auf (Pre: $\alpha = .636$ und $.678$, Trennschärfe = $.35$, vgl.: Bortz 2004). Auch die Schwierigkeitsindizes können mit $.24 - .41$ als ausreichend betrachtet werden. Zu beiden Testzeitpunkten sind die Punkte des ersten Teils des Tests nicht normalverteilt

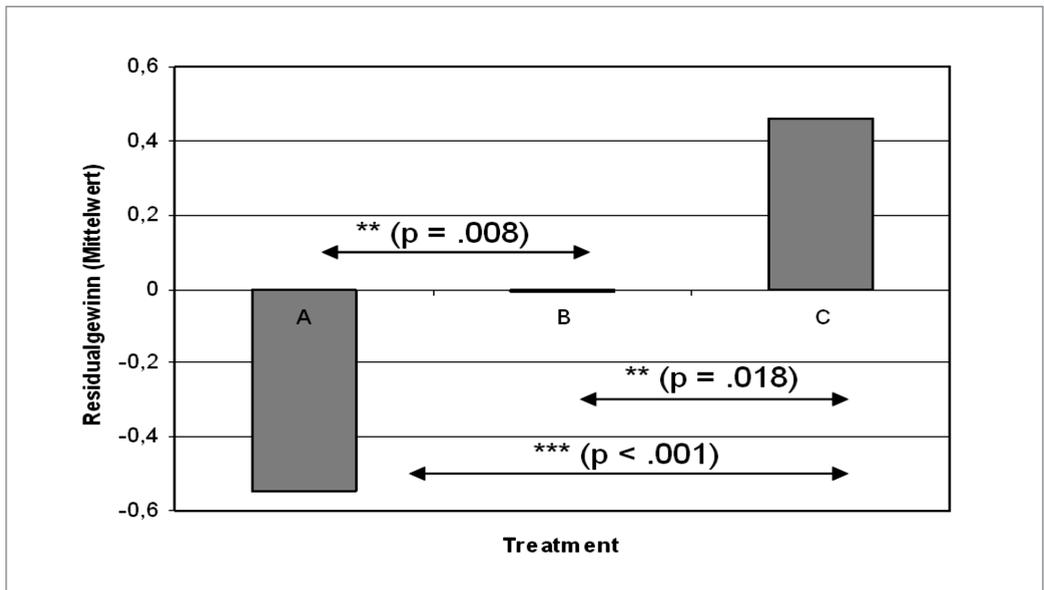


Abb. 5: Residuen der Treatmentgruppen, Hauptstudie

(Pre: K-S-z: 2.079, $p < .001$, Post: K-S-z: 1.911, $p = .001$). Aufgrund dessen konnten die Mittelwerte nicht mittels ANOVA, sondern nur mit dem nichtparametrischen Kruskal-Wallis-Test verglichen werden (vgl.: Brosius 2004, Bortz 2004). Der Kruskal-Wallis-Test gibt an, dass sich die Mittelwerte der Treatmentgruppen zu beiden Testzeitpunkten nicht signifikant von einander unterscheiden (Pre: $p = .279$, Post: $p = .200$). Unter Einbezug der Residuen wird jedoch tendenziell ein Haupteffekt beobachtet ($F(2;164) = 2.294$, $p = .099$, $\eta^2 = .028$), was auf das unterschiedlich hohe Vorwissen der Schüler unter Einbezug der Schulzugehörigkeit zurückzuführen ist. Post-Hoc-Tests zeigen, dass der Mittelwert der Treatmentgruppe C zwar nicht signifikant höher liegt als bei Gruppe A ($p = .182$, $d = .42^3$) und B ($p = .129$, $d = .35$), die mittleren Effekte zeigen jedoch, dass den Interventionen – insbesondere der Intervention in Treatment C – eine praktische Bedeutung zukommt. Signifikanz der Effekte wäre bei einer größeren Stichprobe zu erwarten. Die Intervention in Treatmentgruppe C hat folglich

einen positiven Effekt auf die Lernleistung der SchülerInnen hinsichtlich ihres Wissens im Strukturierungstest. Die SchülerInnen haben mehr Wissen über die Abfolge des Problemlöseprozesses und die beinhalteten Zusammenhänge erlangt, als die SchülerInnen in den Treatmentgruppen A und B.

Auch der zweite Teil des Strukturierungstests weist ausreichende Testgütekriterien auf. Eine Analyse der Mittelwerte im zweiten Teil des Strukturierungstests zum Pre-Zeitpunkt zeigt keinen signifikanten Unterschied zwischen den Treatmentgruppen ($F(2;172) = .308$, $p = .308$, $\eta^2 = .004$), so dass davon ausgegangen werden kann, dass die SchülerInnen gleiches Vorwissen in diesem Bereich aufweisen. Die residualen Mittelwerte unterscheiden sich signifikant von einander ($F(2;165) = 3.279$, $p = .036$, $\eta^2 = .040$). Dieser Haupteffekt ist auf den erhöhten Mittelwert von Treatmentgruppe C gegenüber Treatmentgruppe A ($p = .079$, $d = .42$) und Treatmentgruppe B ($p = .057$, $d = .44$) zurückzuführen.

Auch hier hat die Intervention in der Treatmentgruppe C folglich einen höheren Ef-

³ In der vorliegenden Arbeit werden Post-hoc-Tests immer nach dem Tukey-HSD-Verfahren berechnet.

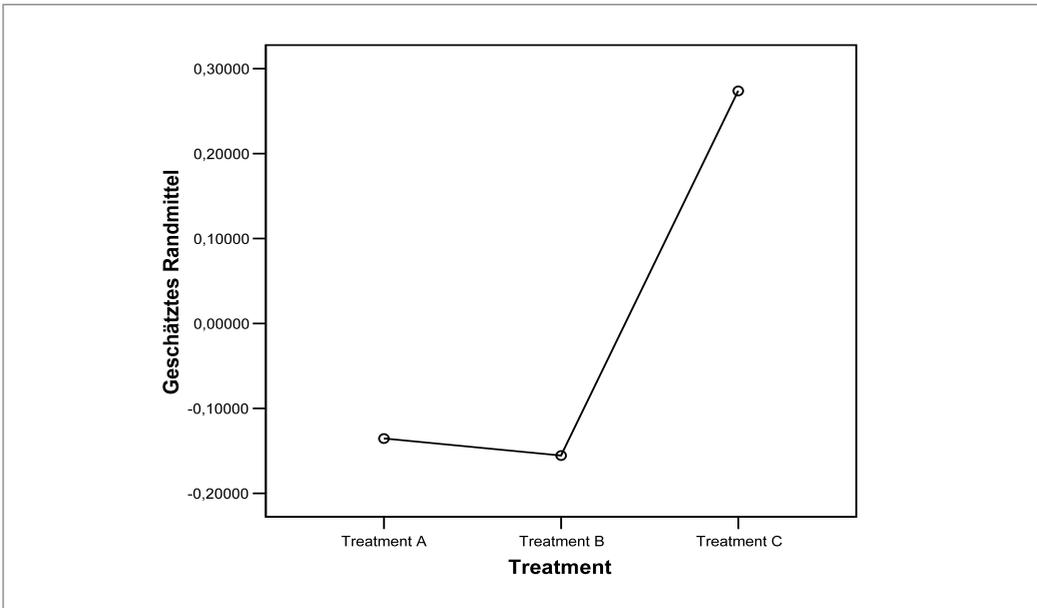


Abb. 6: Residualer Gewinn im zweiten Teil des Strukturierungstests

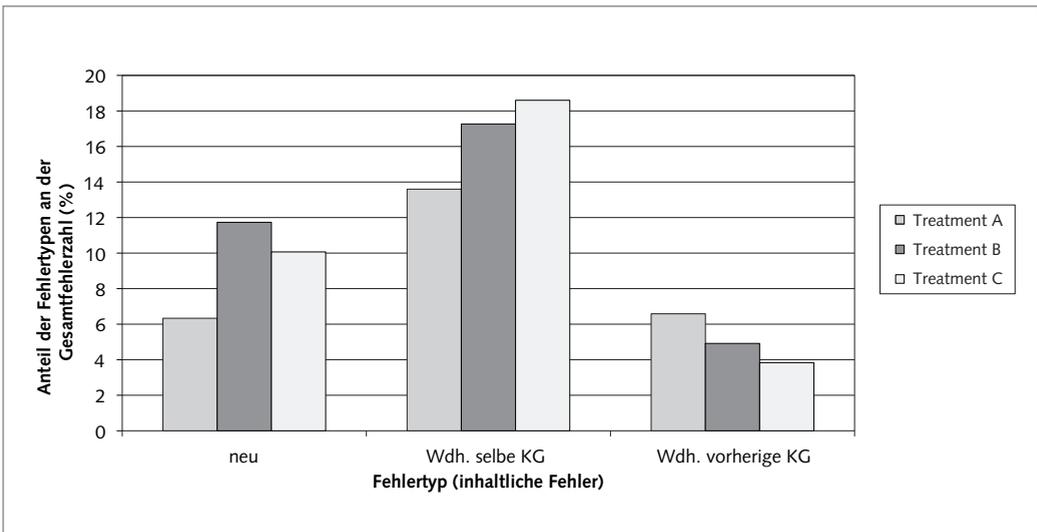


Abb. 7: Prozentuale Darstellung der Typen inhaltlicher Fehler in Bezug auf Treatmentzugehörigkeit

fekt auf den Lernerfolg. Die SchülerInnen des Treatments C können demnach Aussagen besser den einzelnen Schritten des Problemlöseprozesses zuordnen als diejenigen der anderen Treatments. Das durchgeführte Training kann also als erfolgreich angesehen werden. Auch hier hat die Intervention in der

Treatmentgruppe C folglich einen höheren Effekt auf den Lernerfolg. Die SchülerInnen des Treatments C können demnach Aussagen besser den einzelnen Schritten des Problemlöseprozesses zuordnen als diejenigen der anderen Treatments. Das durchgeführte Training kann also als erfolgreich angesehen werden.

6.3 Fehlerkodierung

Die Videos wurden qualitativ ausgewertet, da die Anzahl der Videos in Bezug auf Treatment- und Stundenzugehörigkeit nicht ausreichend für eine quantitative Analyse sind. Die Analyse der Videodaten zeigt, dass eine hohe Fehlerzahl nicht gleichbedeutend mit einer schlecht funktionierenden Gruppenarbeit bzw. mit dem nicht Erreichen des Lernziels sein muss. Stattdessen kann die hohe Fehlerzahl auch darauf zurückgeführt werden, dass die SchülerInnen einer Gruppe z.B. mehr reden als andere oder aber, dass mehr Lösungsmöglichkeiten ausprobiert werden. Es kann demnach nicht von der Fehlerzahl direkt auf den Erfolg und die Qualität der Kleingruppenarbeitsphase geschlossen werden, sondern es bedarf weiterer Analysen. Hierzu werden die vorliegenden Videodaten hinsichtlich inhaltlicher Fehler, manueller Fehler und Strukturierungsfehler genauer betrachtet. Die Auswertung erfolgte mit Hilfe eines standardisierten Manuals, um eine möglichst hohe Objektivität zu erreichen. Darüberhinaus wurde zwischen drei Ratern eine sehr gute Interraterreliabilität ermittelt. Unter inhaltlichen Fehlern wird im Rahmen des Projekts ein Verstoß gegen das in Lernzielen festgehaltene Fachwissen verstanden. Als Beispiel ist die Zuordnung des Glasreinigers zu neutralen Lösungen zu nennen. Manuelle Fehler treten während Experimentalphasen auf, z.B. die gleichzeitige Nutzung von Indikatorstäbchen und Rotkohlsaft. Inhaltliche und manuelle Fehler können jeweils weiter in drei Fehlertypen unterteilt werden: neue Fehler, Wiederholungsfehler in der selben Kleingruppenarbeitsphase und Wiederholungsfehler in einer der folgenden Kleingruppenarbeitsphasen. Unter Wiederholungsfehlern in der selben Kleingruppenarbeitsphase ist dabei ein Fehler zu verstehen, der unverändert wiederholt wird (z.B. bei jeder zu

untersuchenden Lösung werden zwei Indikatoren gleichzeitig eingesetzt). Wird dieser Fehler auch in der folgenden Kleingruppenarbeitsphase wiederholt, so handelt es sich um einen Wiederholungsfehler in einer der folgenden Kleingruppenarbeitsphasen.

Strukturierungsfehler beziehen sich auf die in der Strukturierungshilfe abgebildeten Schritte des naturwissenschaftlichen Problemlöseprozesses. Hier wird zwischen Auslassungs-, Passungs- und Problemlöseprozessfehlern unterschieden. Bei den Auslassungsfehlern werden ein oder zwei Teile der Abfolge Idee, Experiment, Schlussfolgerung nicht durchgeführt (z.B. es wird lediglich ein Experiment durchgeführt, ohne dass zuvor eine Idee oder anschließend eine Schlussfolgerung formuliert wird). Passungsfehler bedeuten, dass die Schritte im durchgeführten Problemlöseprozess nicht zueinander passen. So wird z.B. ein Experiment durchgeführt, mit dem die Idee weder verifiziert noch falsifiziert werden kann: Die Schüler haben die Idee, den pH-Wert der einzelnen Lösungen mit Hilfe eines Indikators zu bestimmen. Im Experiment geben sie zunächst alle Lösungen zusammen und bestimmen dann den pH-Wert. Bei den Problemlöseprozessfehlern wird beobachtet, inwieweit der Problemlöseprozess vollständig durchlaufen wird und ob die einzelnen Schritte richtig oder falsch durchgeführt werden.

Fehlertypen

Bei der Analyse der Typen inhaltlicher Fehler kann beobachtet werden, dass in allen Treatmentgruppen insbesondere in derselben Kleingruppenarbeitsphase Fehler wiederholt werden (Abb. 7). Dies weist darauf hin, dass das angebotene Feedback von den SchülerInnen aller Treatments nicht adäquat genutzt wird⁴. Die prozentuale Fehlerzahl steigt entgegen der Erwartungen von Treatment A zu Treatment C an.

⁴ Beispiel: Den SchülerInnen wird erklärt, dass sie nur wenige Tropfen Indikator einsetzen dürfen, um den pH-Wert der Lösung bestimmen zu können, da ansonsten die Farbe des Indikators die tatsächliche Färbung überdeckt. Im Folgenden geben die SchülerInnen bei der Bestimmung des pH-Werts dennoch (weiterhin) mehrere Milliliter Indikator hinzu, gemäß „viel hilft viel“.

Die Fehler aus den vorhergehenden Kleingruppenarbeitsphasen werden in den nachfolgenden Kleingruppenarbeitsphasen im Gesamtvergleich kaum wiederholt. Dabei ist insbesondere darauf hinzuweisen, dass die prozentuale Anzahl dieser von Treatment A über B zu C abnimmt. Dies ist ein Hinweis darauf, dass der Lehrervortrag und auch das Strukturierungstraining – insbesondere in Kombination – einen positiven Effekt auf die Anzahl dieser Wiederholungsfehler haben.

Die Analyse der Typen manueller Fehler zeigt, dass neue manuelle Fehler prozentual am wenigsten gemacht werden. Dies ist damit zu erklären, dass in den einzelnen Kleingruppenphasen zwar neue Inhalte eingeführt werden, es sich allerdings zumeist um die gleichen oder ähnliche manuelle Arbeitsschritte handelt. Aufgrund dessen kommen am häufigsten Wiederholungsfehler vor. Auch dies weist darauf hin, dass die SchülerInnen das mögliche Feedback nicht adäquat nutzen. Zudem kann in den Videos beobachtet werden, dass die SchülerInnen zwar entsprechendes Feedback erhalten, dieses jedoch nicht annehmen.

Ergänzend ist zu bemerken, dass analog zu den Fehlertypen inhaltlicher Fehler auch bei den manuellen Fehlern die prozentuale Anzahl der Wiederholungsfehler in Bezug auf die vorherigen Kleingruppenarbeitsphasen von Treatment A zu Treatment C abnimmt. Folglich kann auch hier davon ausgegangen werden, dass die Interventionen einen positiven Effekt auf die Zahl manueller Fehler haben.

Die Anzahl möglicher Fehlerinhalte ist nahezu unbegrenzt. Um eine sinnvolle Analyse gewährleisten zu können, wurden mit Hilfe der Daten des Vorgängerprojekts die häufigsten Fehler ermittelt und diese bei der Analyse der neuen Videodaten berücksichtigt. Dennoch wird eine Vielzahl weiterer Fehler beobachtet, die nicht als eigene Kategorien in das Kodierungsmanual aufgenommen werden, sondern als ‚Sonstige‘ zusammengefasst werden. Zu diesen gehören z.B. Aussagen wie:

- „Es kommt immer eine Hauptfarbe bei der Neutralisation heraus“ oder
- „pH-Teststreifen sind ‚Bluthochdruckdinger“.

Bei den inhaltlichen Fehlern werden im Mittel über alle Kleingruppenarbeitsphasen 40 % der Fehler als ‚Sonstige‘ kodiert, bei den manuellen Fehlern 66 %. Bei der Aufbereitung der Darstellung der inhaltlichen und manuellen Fehler wird die Kategorie ‚Sonstige‘ nicht weiter beachtet (Abb. 8 & 9).

Bei den inhaltlichen Fehlern fällt auf, dass insbesondere der Fehler ‚Sinnloses Zusammenmischen‘ vorkommt. Dies ist interessant, da der prozentuale Anteil dieses Fehlers von Treatment A zu Treatment C deutlich abnimmt. Es kann daraus geschlossen werden, dass die Treatmentgruppe C aufgrund des Strukturierungstrainings ihr Vorgehen besser kontrolliert und eher wissenschaftlich orientiert arbeitet, während dies in der Kontrollgruppe nicht der Fall ist. Es handelt sich dort vielmehr um wahlloses Ausprobieren. Diese These wird im Folgenden anhand der Analyse der Strukturierungsfehler weiter untersucht.

Des Weiteren sind auch die Fehler ‚Neutralisation mit Sauerstoff / Stickstoff‘ und ‚Falsche Interpretation Stäbchen‘ in allen Treatments häufig vertreten. Die Videos zeigen, dass dies meist darauf zurückzuführen ist, dass die SchülerInnen häufig ihren Beobachtungen bzw. ihrem selbst durchgeführten Experiment nicht glauben und die erhaltenen Ergebnisse deswegen uminterpretieren.

Auch die Neutralisation einer sauren Lösung mit einer süßen oder neutralen Lösung gehört zu den häufigsten Fehlern in allen Treatmentgruppen. Hierbei handelt es sich um typische Schülervorstellungen (Geisler, 1999), so dass das Auftreten dieser in hoher Zahl nicht überrascht. Unter Einbezug der Ergebnisse des Fachtests kann allerdings bemerkt werden, dass diese Schülervorstellungen insbesondere in Treatmentgruppe C aufgegriffen und fachlich richtige Vorstellungen erreicht werden. Dies zeigt sich darin, dass die SchülerInnen der Treatment-

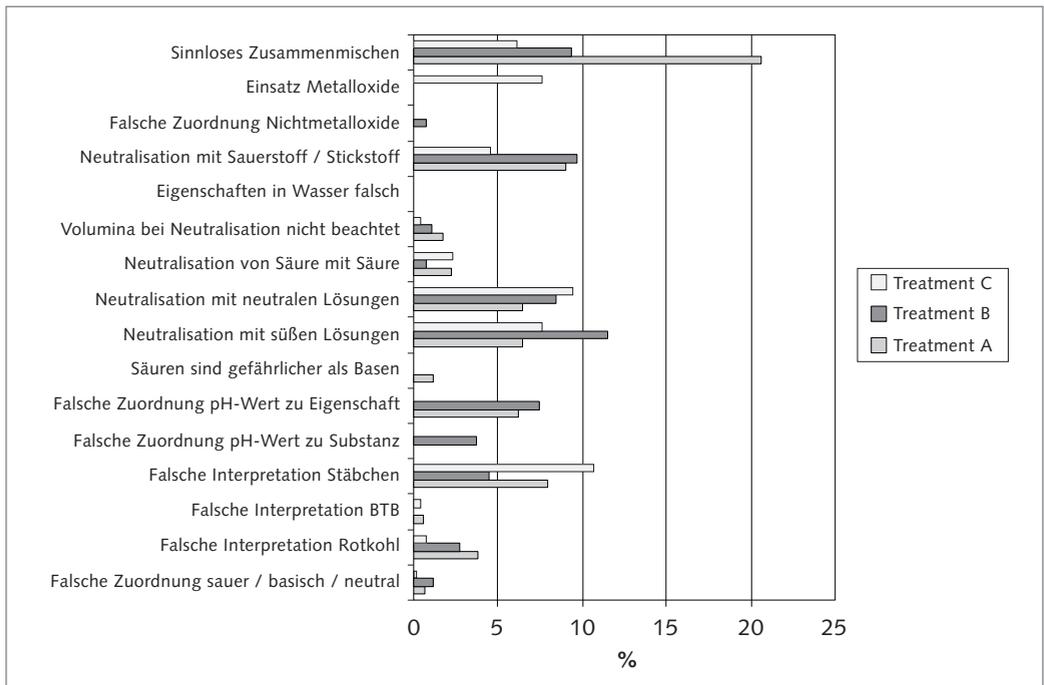


Abb. 8: Inhaltliche Fehler

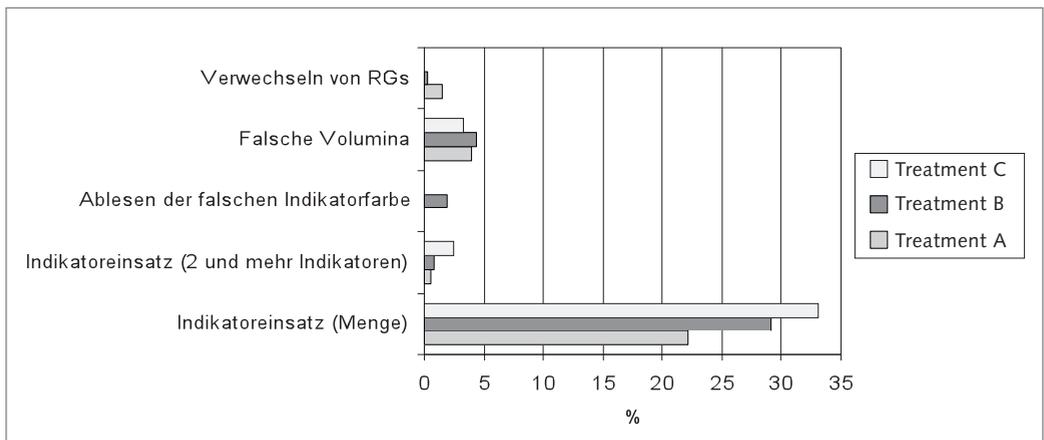


Abb. 9: Manuelle Fehler

gruppe C entsprechende Aufgaben im Fachtest häufiger richtig beantwortet haben.

Die Analyse der Typen manueller Fehler zeigt, dass vor allem der Indikatoreinsatz allen SchülerInnen Schwierigkeiten bereitet (Abb. 9). Dies ist deswegen bemerkenswert, da in mehreren Feedbacksituationen erklärt wird, dass nur wenige Tropfen Indi-

kator eingesetzt werden dürfen. Die SchülerInnen benutzen dennoch zu viel Indikatorlösung. Es lässt sich vermuten, dass sie nicht glauben, dass nur wenige Tropfen für eine sichtbare Reaktion ausreichend sind und aufgrund dessen eine nach ihrem Empfinden ausreichende Menge einsetzen. Auch hier zeigt sich, dass diese Schülervorstellung

fest verankert ist und nicht leicht durch Instruktionen der Lehrenden aufgehoben werden kann.

Die anderen manuellen Fehler kommen nur zu einem geringen Prozentsatz vor.

Strukturierungsfehler

Die Ergebnisse des Strukturierungstests und des Fachtests zeigen, dass insbesondere die Kombination beider Interventionen einen positiven Effekt auf die Schülerleistung hat. Die Videodaten zeigen ähnliche Beobachtungen.

Die Kleingruppenvideos werden zunächst dahingehend kodiert, ob die gestellte Aufgabenstellung richtig, teilweise richtig oder nicht richtig gelöst wird. Einer richtig gelösten Aufgabe wird der Wert 1 zugewiesen, einer teilweise richtig gelösten der Wert 0,5 und einer nicht richtig gelösten der Wert 0. Abbildung 10 zeigt die über alle Kleingruppenphasen gemittelten Werte unter Einbezug der Treatmentzugehörigkeit. Es wird deutlich, dass die Aufgabenstellung häufiger von der Treatmentgruppe C gelöst wird als von den anderen. Dies kann auf ein strukturierteres Vorgehen in dieser Treatmentgruppe

zurückzuführen sein. Um diese These zu überprüfen, werden im Folgenden die anhand der Videodaten erhobenen Strukturierungsfehler näher analysiert.

Bei der Analyse der Videos hinsichtlich der Durchführung von Idee, Experiment und Schlussfolgerung zeigt sich, dass die SchülerInnen der Treatmentgruppe A die meisten Auslassungsfehler – also die Auslassung eines oder mehrerer Schritte im Problemlöseprozess – machen, diejenigen der Treatmentgruppe C am wenigsten (Abb. 11). Dies kann als weiterer Hinweis darauf gewertet werden, dass die Treatmentgruppe C strukturierter vorgeht als die anderen.

Der häufigste Auslassungsfehler aller SchülerInnen ist die Auslassung von Experiment und Schlussfolgerung. In diesen Fällen wird also ausschließlich eine Idee geäußert. Dieses Ergebnis ist zu erwarten, da die SchülerInnen zunächst Ideen äußern, diese mit den Kleingruppenmitgliedern in unterschiedlichem Ausmaß diskutieren, sich dann entweder für eine entscheiden oder auch die anderen gänzlich vergessen. Außerdem werden sie häufig im Unterricht aufgefordert, Ideen erst einmal zu sammeln. Beim

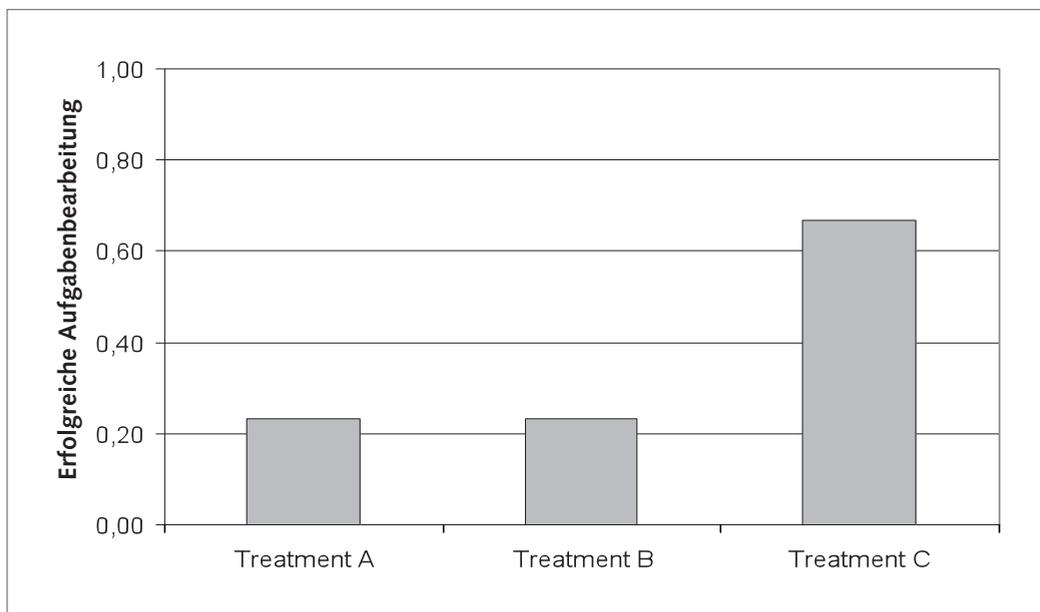


Abb. 10: Erfolgreiche Aufgabenbearbeitung

zweithäufigsten Auslassungsfehler handelt es sich um die Auslassung von Idee und Schlussfolgerung, d.h. es wird nur ein Experiment durchgeführt. Dieser Fehler wird vor allen Dingen von der Treatmentgruppe A gemacht, was wiederum auf ungeplantes Ausprobieren hindeutet. In Treatmentgruppe C kommt dieser Fehlertyp kaum vor. Die Auslassung des Experiments kommt in allen Treatmentgruppen kaum vor. Alle SchülerInnen experimentieren gern. Auch die Auslassung von Idee und Experiment ist

selten zu beobachten, weil die Formulierung einer Schlussfolgerung ohne eine Idee und/oder ein Experiment kaum möglich ist. Diese Situation wird meist dann beobachtet, wenn Lernende Vorwissen einbringen und von der Richtigkeit dieses Wissens überzeugt sind. Sie formulieren dann häufig Behauptungen im Sinne einer Schlussfolgerung, ohne diese weiter zu begründen oder zu überprüfen. Auch die Passungsfehler können einen Hinweis auf das strukturierte und erfolgreiche Vorgehen der SchülerInnen geben (Abb. 12).

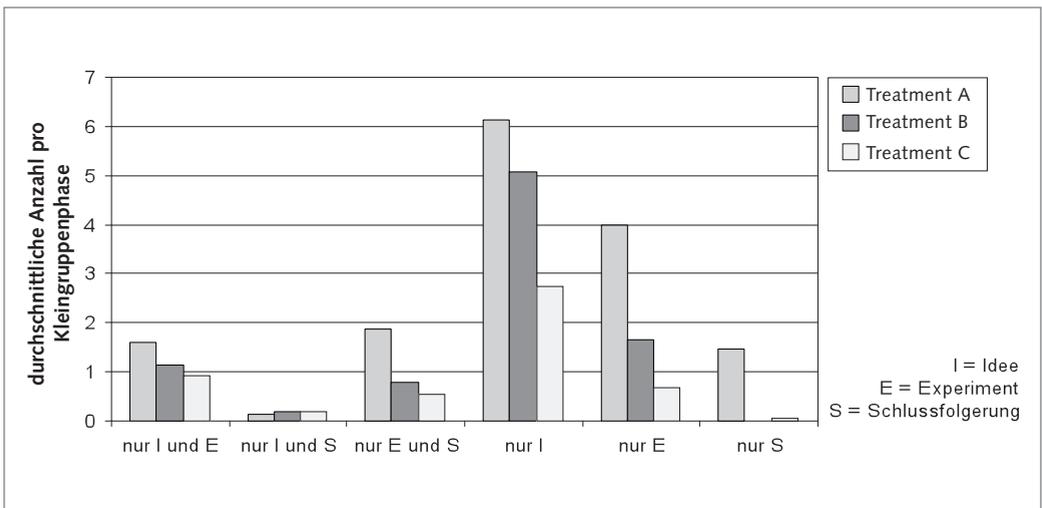


Abb. 11: Auslassungsfehler

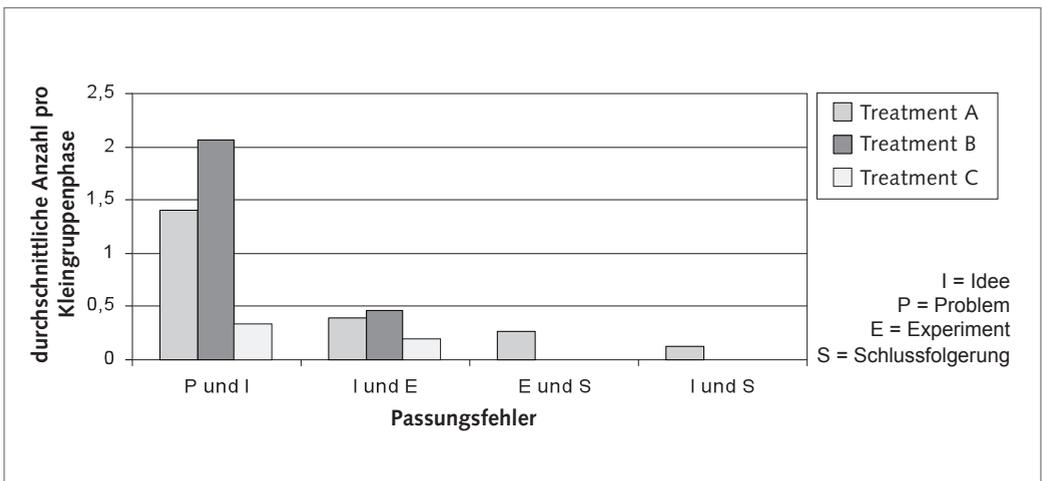


Abb. 12: Passungsfehler

Es zeigt sich, dass hier die Treatmentgruppen A und B die meisten Fehler machen. Dies kann damit erklärt werden, dass die SchülerInnen zum einen unstrukturiert vorgehen und zum anderen nicht verstanden haben, was zu tun ist, um zu einem sinnvollen Ergebnis zu gelangen.

Häufig kann beobachtet werden, dass keine Passung zwischen Problem und Idee vorliegt. Die SchülerInnen haben also entweder das Problem nicht verstanden oder sind nicht motiviert, theoriegeleitet zu arbeiten, sondern ziehen vor, alles zu mischen. Es ist interessant zu beobachten, dass dieser Fehler am häufigsten von der Treatmentgruppe B gemacht wird. Hier scheint der Lehrervortrag keinen Einfluss zu haben, während das Strukturierungstraining in Kombination mit dem Lehrervortrag einen großen Effekt aufweist. Die übrigen Passungsfehler werden kaum beobachtet werden.

Abbildung 13 gibt die Anzahl vollständiger Problemlöseprozesse pro Kleingruppenarbeitsphase an und zudem wird differenziert, ob die einzelnen Schritte des Problemlöseprozesses (Idee, Experiment, Schlussfolgerung) richtig oder falsch (f) durchgeführt werden.

Die Ergebnisse zeigen, dass auch hier die Treatmentgruppen B und C vor der Treatmentgruppe A liegen, wiederum ein Hinweis, dass die SchülerInnen der Treatmentgruppe A nicht strukturiert vorgehen

Es ist allerdings hervorzuheben, dass die richtige Durchführung des vollständigen Prozesses (I/E/S) insbesondere in der Treatmentgruppe C gelingt. Die Treatmentgruppe B und vor allem die Treatmentgruppe A liegen deutlich hinter dieser zurück. Demnach kann davon ausgegangen werden, dass die Kombination der beiden Interventionen einen positiven Effekt auf das strukturierte und richtige Vorgehen der SchülerInnen hat, während der Lehrervortrag allein lediglich ein strukturierteres Vorgehen fördert. Gestützt wird diese These durch die Beobachtung, dass die SchülerInnen der Treatmentgruppe B bei den beobachteten Problemlöseprozessen deutlich häufiger eine falsche Schlussfolgerung formulieren als die anderen (I/E/fS, I/fE/fS, fI/E/fS, fI/fE/fS). Hierzu kann folgendes Fallbeispiel angeführt werden: Die SchülerInnen haben die Aufgabe den pH-Wert von hautneutraler Seifenlösung zu bestimmen. Die SchülerInnen haben die richtige Idee, den pH-Wert mit

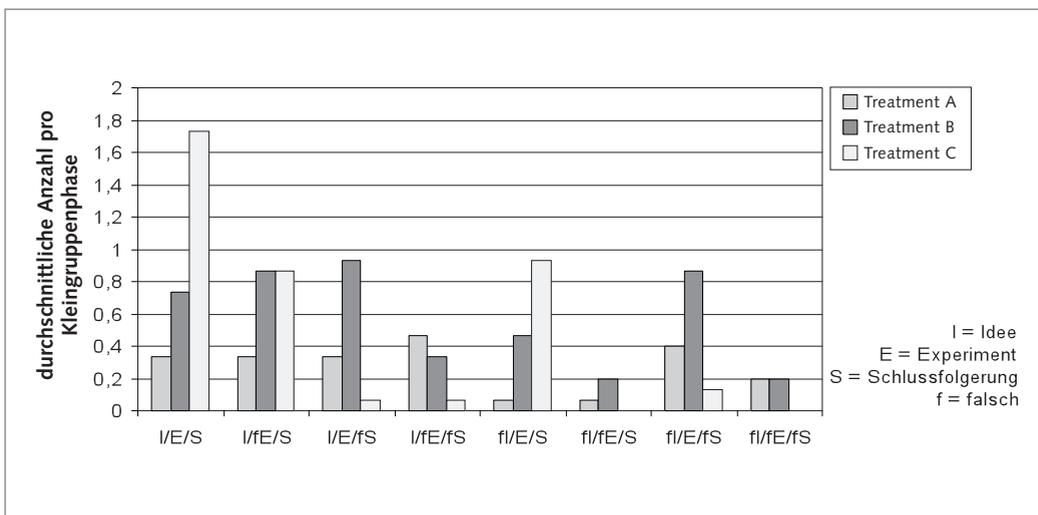


Abb. 13: Problemlöseprozess

Rotkohlsaft zu bestimmen und führen das entsprechende Experiment richtig durch. Sie beobachten, dass sich die Lösung rot färbt. Gemäß der ihnen vorliegenden Skala des Rotkohllindikators würde dies einen sauren pH-Wert bedeuten. Die SchülerInnen vertrauen ihren eigenen Beobachtungen jedoch nicht, da in dem Wort „pH-hautneutral“ der Begriff „neutral“ vorkommt. Sie haben schon einmal gehört, dass neutral ein pH-Wert von 7 bedeutet. Aufgrund dessen formulieren sie die falsche Schlussfolgerung: pH-hautneutrale Seifenlösung hat einen pH-Wert von 7 (→ I/E/fs)

Es kann gefolgert werden, dass sowohl der Lehrervortrag als auch die Kombination von Lehrervortrag und Strukturierungstraining einen positiven Effekt auf das strukturierte Vorgehen der SchülerInnen haben, aber die Kombination beider Maßnahmen zu stärkerer Strukturierung des Vorgehens führt und diese SchülerInnen auch häufiger das Ziel der Kleingruppenarbeitsphase erreichen. Die Kombination der Interventionen ist folglich der Intervention ‚Lehrervortrag‘ überlegen.

7 Zusammenfassung

Zusammenfassend ist festzuhalten, dass die SchülerInnen der Treatmentgruppe C die gestellten Aufgaben häufiger richtig beantworteten, manuelle und inhaltliche Fehler weniger häufig wiederholen und strukturierter vorgehen.

Einen ersten Hinweis auf die Wirksamkeit des Strukturierungstrainings in Bezug auf das strukturierte experimentelle Arbeiten der SchülerInnen findet sich bei den inhaltlichen Fehlern. Die SchülerInnen der Treatmentgruppe A äußern deutlich häufiger die Idee, das Problem durch ‚Zusammenschütten‘ zu lösen, als diejenigen der beiden anderen Treatmentgruppen. Die Zahl dieser Fehler nimmt von Treatment A über B zu C ab. Ähnliche Ergebnisse zeigt die Analyse der Strukturierungsfehler. Bei den SchülerInnen der Treatmentgruppe A kommen deutlich

häufiger allein stehende falsche Experimente vor. Auch hierbei handelt es sich häufig um ungeplantes Zusammengeben von Chemikalien. Zudem gehen die SchülerInnen der Treatmentgruppe C im Problemlöseprozess deutlich häufiger richtig strukturiert vor als die übrigen. Diese Beobachtungen werden durch die Ergebnisse des Strukturierungstests unterstützt. Die SchülerInnen haben also nicht nur Wissen über einen strukturierten Problemlöseprozess erworben, sondern wenden dieses Wissen auch in konkreten Situationen an.

Insgesamt zeigt die Analyse des Fachtests, dass die Lernumgebung einen positiven Effekt auf die Lern- und Behaltensleistung der SchülerInnen hat. Im Pre-Post-Vergleich zeigt sich zudem, dass die SchülerInnen der Treatmentgruppe C signifikant besser abschneiden als die anderen, während gleichzeitig diejenigen der Treatmentgruppe B auch signifikant besser abschneiden als in Gruppe A. Der Lehrervortrag hat also einen positiven Effekt auf die Schülerleistung, der durch die Kombination mit dem Strukturierungstraining vergrößert wird. Der Strukturierungstest zeigt darüber hinaus an, dass die Treatmentgruppe C besser abschneidet als die anderen. Die beiden anderen Gruppen unterscheiden sich nicht. Die SchülerInnen des Treatments C besitzen mehr Wissen über den Problemlöseprozess als die anderen und wenden dieses Wissen auch an, wie die Videodaten zeigen. Sie gehen häufiger richtig strukturiert vor als die SchülerInnen der beiden anderen Gruppen. Aus den besseren Testergebnissen und der höheren Wahrscheinlichkeit der richtigen Aufgabenbearbeitung und -beantwortung in den Kleingruppenarbeitsphasen kann gefolgert werden, dass das richtig strukturierte Vorgehen einen positiven Effekt auf die Schülerleistung hat.

Die Forschungsfrage lässt sich demnach insgesamt positiv beantworten: Mit Hilfe des durchgeführten Strukturierungstrainings kann der Umgang mit den strukturierenden Lernprozesshilfen gefördert und somit der Lernerfolg der SchülerInnen gesteigert wer-

den. Dabei ist zu bemerken, dass die Interventionen einzeln betrachtet bereits einen positiven Effekt auf die Schülerleistung haben (vgl. Pilotstudie), dieser Effekt allerdings durch die Kombination der beiden Interventionen deutlich gesteigert werden kann.

Die in der Studie entwickelten Hilfen und unterrichtlichen Inhalte können sicherlich auch im ‚normalen‘ Unterricht eingesetzt werden. Da es sich bei der vorgestellten Studie allerdings um eine experimentelle Laborstudie handelt, können keine direkten Aussagen über den Erfolg der Intervention im Schulalltag getätigt werden. Aufgrund der gezeigten positiven Wirkung der Intervention auf die Lernleistung, das naturwissenschaftliche Problemlösen und die Selbsttätigkeit der SchülerInnen, kann jedoch vermutet werden, dass der Einsatz der entwickelten Hilfen und der unterrichtlichen Inhalte in der Schule durchaus einen positiven Effekt auf die o.g. Faktoren hat.

Wir danken der Deutschen Forschungsgemeinschaft für ihre finanzielle Unterstützung (SU 187/6-1 bis 6-3) und den SchülerInnen für ihre Teilnahme an der Untersuchung.

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, 665-701.
- Abell, S.K. & Lederman, N.G. (2007). *Handbook of Research on Science Education*. Mahwah, NJ: Lawrence Erlbaum.
- Akerson, V.L. & Hanuscin, D.L. (2007). Teaching Nature of Science through Inquiry: Results of a 3-Year Professional Development Program. *Journal of Research in Science Teaching*, 44, 653-680.
- Akerson, V.L. & Volrich, M.L. (2006). Teaching Nature of Science Explicitly in a First-Grade Internship Setting. *Journal of Research in Science Teaching*, 43, 337-394.
- American Association for the Advancement of Science (AAAS), Project 2061 (1990). *Science for All Americans*. New York: Oxford University Press.
- American Association for the Advancement of Science (AAAS), Project 2061 (1993). *Benchmarks for Science Literacy*. New York: Oxford University Press.
- American Association for the Advancement of Science (AAAS), Project 2061 (2001). *Atlas of Science Literacy*. New York: Oxford University Press.
- Anderson, J.R., Corbett, A.T., Koedinger, K. R. & Pelletier, R. (1995). Cognitive tutors: Lessons Learned. *Journal of the Learning Sciences*, 4, 167-207.
- Atkin, J.M. & Black, P. (2007). History of Science Curriculum Reform in the United States and the United Kingdom. In S. K. Abell & N.G. Lederman (Eds.), *Handbook of Research on Science Education* (pp.781-806). Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Bangert-Drowns, R.L., Kulik, C.-L., Kulik, J.A. & Morgan, M.T. (1991). The Instructional Effect of Feedback in Test-Like Events. *Review of Educational Research*, 61, 213-238.
- Barrow, L. H. (2006). A Brief History of Inquiry: From Dewey to Standards. *Journal of Science Teacher Education*, 17, 265-278.
- Bäurle, W., Barmeier, M., Ciprina, H.J., Hell, K., Hoppe, B., Maiworm et al. (2004). *Das Methodenbeft für Naturwissenschaften. Strategie Sek. I*. Stuttgart: Klett Verlag.
- Bell, R.L., Blair, L.M., Crawford, B.A. & Lederman, N.G. (2003). Just Do It? Impact of a Science Apprenticeship Program on High School Students' Understandings of the Nature of Science and Scientific Inquiry. *Journal of Research in Science Teaching*, 40, 487-509.

- Bortz, J. (2004). *Statistik für Sozialwissenschaftler*. Berlin: Springer Verlag.
- Brosius, F. (2004). *SPSS 12*. Bonn: mitp-Verlag.
- Brown, P.L., Abell, S.A., Demir, A. & Schmidt, F.J. (2006). College Science Teachers' View of Classroom Inquiry. *Science Education*, 90, 784-802.
- Butler, D.L. & Winne, P.H. (1995). Feedback and Self-Regulated Learning: A Theoretical Synthesis. *Review of Educational Research*, 65, 245-281.
- Bybee, R.W. (1997). Toward an Understanding of Scientific Literacy. In W. Gräber & C. Bolte (Hrsg.), *Scientific Literacy. An International Symposium* (pp.37-68). IPN-Schriftenreihe Nr. 154. Kiel: Institut für die Pädagogik der Naturwissenschaften (IPN).
- Dörner, D. (1976). *Problemlösen als Informationsverarbeitung*. Stuttgart: Kohlhammer.
- Erdfelder, E., Faul, F. & Buchner, A. (1996). GPOWER: A general power analysis program. *Behavior Research Methods, Instruments, & Computers*, 28, 1-11.
- Farquhar, J.D. (1995). *A summary of research with the Console-Operations Tutor: LOADER*. <http://www2.gsu.edu/~wwwitr/research/farquhar1995.htm> (zuletzt aufgesucht am 08.01.09).
- Flick, L.B. & Lederman, N. (2004). *Scientific Inquiry and Nature of Science. Implications for Teaching, Learning, and Teacher Education*. Dordrecht: Kluwer Academic Press.
- Gallagher-Bolos, J.A. & Smithenry, D.W. (2004). *Teaching Inquiry-Based Chemistry. Creating Student-Led Scientific Communities*. Portsmouth, NH: Heinemann.
- Geisler, A. (1999). *Chemielernen als Änderung von Schülervorstellungen am Beispiel Säure-Base*. Dissertation an der Universität-Gesamthochschule Essen.
- Glück, G. (1999). Zeitgeist und Fehlertheorie (1921 – 1939). Meister Weimer und sein Schüler Kießling. In W. Althof (Hrsg.), *Fehlerwelten. Vom Feblemachen und Lernen aus Fehlern* (S.169-187). Opladen: Leske + Budrich.
- Hammerer, F. (2001). Der Fehler – eine pädagogische Schlüsselssituation und Herausforderung. *Erziehung und Unterricht*, 151(1-2), 37-50.
- Hattie, J. & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research*, 77, 81-112.
- Hübinger, R. & Sumfleth, E. (2006). *Materialien für den naturwissenschaftlichen Unterricht für die Klassen 5/6. „Mein Körper und ich auf Weltreise“*. Berlin: Schering Stiftung.
- Jacobs, B. (1998). Die Bedeutung des Feedbacks im normalen Schulunterricht – die Metaanalyse von Lysakowsky und Walberg (1982). <http://www.phil.uni-sb.de/~jakobs/www/artikel/feedback/lysakow.htm> (zuletzt aufgesucht am 08.01.09).
- Jacobs, B. (2002). Aufgaben stellen und Feedback geben. <http://www.phil.uni-sb.de/~jakobs/www/artikel/feedback/feedback.pdf> (zuletzt aufgesucht am 08.01.09).
- Kahl, R. (2002). Der Fehler ist das Salz des Lernens. *Das Magazin*, 6(1), 9-11.
- Klahr, D. (2000). *Exploring Science. The cognition and development of discovery processes*. Cambridge, MA: MIT Press.
- Klieme, E., Avenarius, H., Blum, W., Döbrich, P., Gruber, H. Prenzel, M. et al. (Hrsg.) (2003). *Expertise zur Entwicklung nationaler Bildungsstandards*. Berlin: BMBF.
- Kulhavy, R.W., Stock, W.A., Hancock, T.E., Swindell, L.K. & Hammrich, P.L. (1990). Written feedback: Response certitude and durability. *Contemporary Educational Psychology*, 15, 319-332.
- Kultusministerkonferenz (KMK) (2005). *Bildungsstandards im Fach Chemie für den Mittleren Schulabschluss*. Neuwied: Luchterhand.
- Lederman, J. & Lederman, N.G. (2006). *Discovering Inquiry and Nature of Science: Definitions, Distinctions, Teaching, and Assessment*. Folien zum Vortrag bei der Forschergruppe und dem Graduiertenkolleg 'Naturwissenschaftlicher Unterricht' in Essen.
- Lederman, N.G. (2007). Nature of Science: Past, Present, and Future. In S.K. Abell & N.G. Lederman (Hrsg.), *Handbook of Research on Science Education* (pp.830-879). Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Lysakowski, R.S. & Walberg, H.J. (1982). Instructional effects of cues, participation, and corrective feedback. A quantitative synthesis. *American Educational Research Journal*, 19, 559-578.
- Martin-Hauser, L. (2002). Defining Inquiry. *The Science Teacher*, 69(2), 34-37.
- Mathan, S.A. & Koedinger, K.R. (2005). Fostering the Intelligent Novice: Learning From Errors With Metacognitive Tutoring. *Educational Psychologist*, 40, 257-265.
- McComas, W.F. (Hrsg.) (1998). *The Nature of Science in Science Education. Rationales and Strategies*. Dordrecht: Kluwer Academic Press.
- Merrill, D.M. (1987). A lesson based on the component display theory. In C.M. Regeluth (Hrsg.), *Instructional Theories in Action* (pp.201-244). Hillsdale, NJ: Lawrence Erlbaum Associates.

- Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen (Hrsg.) (2004). *Chemie. Sekundarstufe I. Gymnasium. Richtlinien und Lehrpläne*. Frechen: Ritterbach.
- Minstrell, J. & van Zee, E.H. (Hrsg.) (2000). *Inquiring Into Inquiry Learning and Teaching in Science*. Washington, DC: American Association for the Advancement of Science (AAAS).
- Müller, A. & Nieswandt, M. (1999). *Arbeitsleitfaden zu Modul 3: „Aus Fehlern lernen“*. BLK-Programm: *Steigerung der Effizienz des mathematisch-naturwissenschaftlichen Unterrichtes*. Kiel: Institut für die Pädagogik der Naturwissenschaften an der Universität Kiel (IPN).
- Narciss, S. (2004). *Informatives tutorielles Feedback. Ableitung und empirische Überprüfung von Entwicklungs- und Evaluationsprinzipien auf der Basis instruktionspsychologischer Erkenntnisse*. Dresden: Technische Universität Dresden. Fakultät Mathematik und Naturwissenschaften.
- Narciss, S & Huth, H.M. (2004). How to design informative tutoring feedback for multimedia learning. In R. Brüncken & D. Leutner (Hrsg.), *Instructional Design for Multimedia Learning* (S.181-195). Münster: Waxmann.
- National Research Council (NRC) (2000). *Inquiry and the National Science Education Standards. A Guide for Teaching and Learning*. Washington, D.C.: National Academic Press.
- Newell, A. & Simon, H.A. (1972). *Human Problem Solving*. Englewood Cliffs, New Jersey: Prentice-Hall.
- Niedderer, H. (1996). Überblick über Lernprozessstudien in Physik. In R. Duit & C. von Rhöneck (Hrsg.), *Lernen in den Naturwissenschaften* (S.119-144). Kiel: Leibnitz-Institut für die Pädagogik der Naturwissenschaften an der Universität Kiel (IPN).
- Osborne, J., Ratcliffe, M., Collins, S., Millar, R. & Duschl, R. (2003). What 'ideas-about-science' should be taught in school science? A Delphi study of the expert community. *Journal of Research in Science Teaching*, 40, 692-720.
- Oser, F., Hascher, T. & Spychiger, M. (1999a). Lernen aus Fehlern. Zur Psychologie des „negativen“ Wissens. In W. Althof (Hrsg.), *Fehlerwelten: Vom Fehlermachen und Lernen aus Fehlern. Beiträge und Nachträge zu einem interdisziplinären Symposium aus Anlaß des 60. Geburtstags von Fritz Oser* (S.11-42). Opladen: Leske + Budrich.
- Oser, F. & Spychiger, M. (2000). Lernen aus Fehlern als Beitrag zum Lebenslangen Lernen. In F. Achtenhagen & W. Lempert (Hrsg.), *Lebenslanges Lernen im Beruf Seine Grundlegung im Kindes- und Jugendalter (IV). Formen und Inhalte von Lernprozessen* (S.101-122). Opladen: Leske + Budrich.
- Oser, F. & Spychiger, M. (2005). *Lernen ist schmerzhaft. Zur Theorie des Negativen Wissens und zur Praxis der Fehlerkultur*. Weinheim: Beltz Verlag.
- Oser, F., Spychiger, M., Hascher, T. & Mahler, F. (1997a). *Die Fehlerkulturschule. Entwicklung der Fehlerkultur als Projekt im Rahmen von Schulentwicklung*. Schriftenreihe zum Projekt „Lernen Menschen aus Fehlern? Zur Entwicklung einer Fehlerkultur in der Schule“, Nr. 3. Freiburg, Schweiz: Pädagogisches Institut der Universität Freiburg.
- Oser, F., Spychiger, M., Mahler, F. & Hascher, T. (1997b). *1. wissenschaftlicher Zwischenbericht des Forschungsprojekts „Lernen Menschen aus Fehlern? Zur Entwicklung einer Fehlerkultur in der Schule*. Universität Freiburg, Schweiz: Lehrstuhl für Pädagogik und pädagogische Psychologie.
- Oser, F., Spychiger, M., Mahler, F., Hascher, T., Gut, K., Buehler, U. et al. (1998). *2. wissenschaftlicher Zwischenbericht des Forschungsprojekts „Lernen Menschen aus Fehlern? Zur Entwicklung einer Fehlerkultur in der Schule*. Universität Freiburg, Schweiz: Lehrstuhl für Pädagogik und pädagogische Psychologie.
- Oser, F., Spychiger, M., Mahler, F. & Reber, S. (1999b). *3. wissenschaftlicher Zwischenbericht des Forschungsprojekts „Lernen Menschen aus Fehlern? Zur Entwicklung einer Fehlerkultur in der Schule*. Universität Freiburg, Schweiz: Lehrstuhl für Pädagogik und pädagogische Psychologie.
- Pommer, Mona (2004). *Informatives Feedback. Wirkung auf Lernerfolg und Motivation bei computergestütztem Training sprachrezeptiver Fähigkeiten*. Münster: Waxmann.
- Pridemore, D.R. & Klein, J.D. (1995). Control of Practice and Level of Feedback in Computer-Based Instruction. *Contemporary Educational Psychology*, 20, 444-450.
- Roberts, D.R. (2007). Scientific Literacy / Science Literacy. In S.K. Abell & N.G. Lederman (Eds.), *Handbook of Research on Science Education* (pp.729-779). Mahwah, New Jersey. Lawrence Erlbaum Associates.
- Rollett, B. (1999). Auf dem Weg zu einer Fehlerkultur. Anmerkungen zur Fehlertheorie von Fritz Oser. In W. Althof (Hrsg.), *Fehlerwelten: Vom Fehlermachen und Lernen aus Fehlern* (S.71-87). Opladen: Leske + Budrich.
- Rumann, S. (2005). Kooperatives Arbeiten im Chemieunterricht. Entwicklung und Evaluation einer Interventionsstudie zur Säure-Base-Thematik. Aus der Reihe H. Niedderer, H. Fischler & E. Sumfleth (Hrsg.), *Studien zum Physik- und Chemielernen*. Band 46. Berlin: Logos Verlag.

- Schwartz, R.S., Lederman, N.G. & Crawford, B.A. (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, 610-645.
- Spychiger, M., Oser, F., Hascher, T., Gut, K. & Büeler, U. (2000). Auf dem Weg zu einer Fehlerkultur in der Schule – Serie 7: Fehler beim Namen nennen. *Die neue Schulpraxis*, 70(4), 12-14.
- Spychiger, M., Oser, F., Hascher, T. & Mahler, F. (1997). *Untersuchung und Veränderung von Fehlerkultur in der Schule. Der Situationsansatz. Schriftenreihe zum Projekt „Lernen Menschen aus Fehlern? Zur Entwicklung einer Fehlerkultur in der Schule“*. Nr. 2. Freiburg, Schweiz: Pädagogisches Institut der Universität Freiburg.
- Spychiger, M., Oser, F., Hascher, T. & Mahler, F. (1999). Entwicklung einer Fehlerkultur in der Schule. In W. Althof (Hrsg.), *Fehlerwelten. Vom Fehlermachen und Lernen aus Fehlern* (S.43-70). Opladen: Leske + Budrich.
- Stäudel, L., Werber, B., & Freiman, T., (2002). *Lernbox Naturwissenschaften. Verstehen und anwenden*. Seelze. Friedrich Verlag.
- Walpuski, M. (2006). Optimierung von experimenteller Kleingruppenarbeit durch Strukturierungshilfen und Feedback. Aus der Reihe H. Niedderer, H. Fischler & E. Sumfleth (Hrsg.), *Studien zum Physik- und Chemielernen*. Band 49. Berlin: Logos Verlag.
- Walpuski, M. & Sumfleth, E. (2007). Strukturierungshilfen und Feedback zur Unterstützung experimenteller Kleingruppenarbeit im Chemieunterricht. *Zeitschrift für Didaktik der Naturwissenschaften*, 13, 181-198.
- Wirth, J. & Leutner, D. (2006). Selbstregulation beim Lernen in interaktiven Lernumgebungen. In H. Mandel & H.F. Friedrich (Hrsg.), *Handbuch Lernstrategien* (S.172-184). Göttingen: Hogrefe.
- Weimer, H. (1925). *Psychologie der Fehler*. Leipzig: Klinkhardt.

Kontakt

Prof. Dr. Elke Sumfleth,
Universität Duisburg-Essen
Didaktik der Chemie
Schützenbahn 70
45127 Essen
elke.sumfleth@uni-due.de

Autoreninformation

Dr. Isabel Wahser war Mitglied des Graduiertenkollegs und der Forschergruppe nwu an der Universität Duisburg-Essen. Dr. Isabel Wahser ist zur Zeit Referendarin am Studienseminar Duisburg – Gymnasium/Gesamtschule mit den Fächern Chemie, Biologie und Naturwissenschaften. Dr. Elke Sumfleth ist Professorin für Didaktik der Chemie an der Universität Duisburg-Essen und Sprecherin des DFG-Graduiertenkollegs „Naturwissenschaftlicher Unterricht“ an der Universität Duisburg-Essen. Ihre Forschungsschwerpunkte sind: Empirische Lehr-Lern-Forschung in Chemie, Entwicklung und Evaluation von Materialien für den Chemieunterricht und Kompetenzentwicklung durch Aufgaben im Chemieunterricht.