

Der Einfluss von Schüler- und Demonstrationsexperimenten auf den Lernzuwachs in Physik

Jan Winkelmann*, Roger Erb*

*Goethe-Universität Frankfurt, Institut für Didaktik der Physik
winkelmann@physik.uni-frankfurt.de, roger.erb@physik.uni-frankfurt.de
(Eingegangen: 05.10.2017; Angenommen: 26.04.2018)

Kurzfassung

Die aktuelle Forschungslage zum Einfluss des Experimentierens in den Naturwissenschaften auf die fachlichen Leistungen von Schülerinnen und Schülern zeichnet ein uneinheitliches Bild: Manche Studien sehen Schülerexperimente im Vorteil, andere beschreiben bessere Lernerfolge durch Demonstrationsexperimente. Die vorliegenden Ergebnisse sind nur schwer miteinander zu vergleichen, da sich nicht nur das Design der Studien stark voneinander unterscheidet, sondern auch Variablen wie die realisierte Einbettung in den Unterricht, von denen ebenfalls ein Einfluss auf den Lernerfolg vermutet werden kann, oft nur unzureichend kontrolliert wurden. Dies führt dazu, dass in jüngerer Forschung einzelne Variationen im Unterrichtsablauf nur noch selten verglichen werden. Demgegenüber steht die Forderung, dass sich die Erkenntnisse, die in Studien zur empirischen Unterrichtsforschung gewonnen werden, letztlich auch auf das Unterrichtshandeln der Lehrkräfte auswirken müssten.

In der vorliegenden Vergleichsstudie wurde daher der Einfluss verschiedener Arten des Experimentierens im Physikunterricht auf den Lernzuwachs der Schülerinnen und Schüler (n = 858) untersucht. Dabei war das Ziel zum einen die Einbettung in den Unterricht weitgehend zu kontrollieren und zum anderen die Lehrkräfte selbst als unterrichtswirksamen Faktor explizit einzubinden. Die hierfür eigens erstellten Treatments haben die geometrische Optik zum Gegenstand, bei der Studie selbst handelt es sich um eine quasi-experimentelle Feldstudie mit einem Pre-Posttest-Design.

Es konnte gezeigt werden, dass die Art des Experimentierens keinen signifikanten Einfluss auf den Lernerfolg hat. Allerdings zeigte sich, dass es wichtig ist, welche Lehrkraft mit welcher Art des Experimentierens unterrichtet. Diese Wechselwirkung führte zu (kleinen) Effekten zwischen den Vergleichsgruppen.

Abstract

The state of research on the influence of experimenting in science on students' achievement draws a heterogeneous picture: Some studies favor doing practical work, whereas others commend teacher demonstrations. It is difficult to compare the available results, due not only in part to the greatly differing study designs, but also to the insufficient provision of the control of variables as well as the realization of the embedded lessons. This means, on the one hand, the contemporary research now rarely compares single variations of normal teaching tasks. In contrast, there is the requirement that the findings of empirical education research should be implemented in teachers' action in classrooms.

For this purpose, we tested within the framework of a comparative study (n = 858) if different ways of experimenting in school science have an influence on the growth of students' knowledge. The aim was to control the integration of the intervention into the lessons on one hand and to integrate the teacher as an effective factor of learning success into the analyses on the other hand. The purpose-built treatments deal with geometrical optics in physics, the study itself is a quasi-experimental field study with a pre-posttest design.

The study has shown that the way of doing experiments in school does not have a significant impact on learning success. However, the study showed that there is interdependency between the teacher and the particular kind of experimenting, which explains the (small) differences between our comparison-groups.

1. Einleitung

Das Experimentieren in den Naturwissenschaften ist eine etablierte Methode, um Antworten auf offene Fragen zu erhalten, die in Bezug zu einem theoretischen Hintergrund stehen. Diese Methode findet sich auch im Physikunterricht wieder, in dem die angesprochenen Fragestellungen neu für die Schülerinnen und Schüler sind, obwohl sie in der Wissenschaft bereits gelöst sind. Neben der Entwicklung des fachlichen Wissens zielen Experimente darüber hinaus auch darauf ab, naturwissenschaftliches Arbeiten zu vermitteln und das Interesse sowie die Motivation der Schülerinnen und Schüler zu steigern (Welzel et al., 1998).

Neuere Studien zur Rolle des Experimentierens im naturwissenschaftlichen Unterricht sind daher auch eng mit etwaigen Auswirkungen auf das Interesse, die Motivation und Haltungen gegenüber den Naturwissenschaften sowie mit Auswirkungen auf fachliche Leistungen von Schülerinnen und Schülern verknüpft (Osborne, Simon & Collins, 2003; Millar, 2010; Potvin & Hasni, 2014).

Diese Verbindung drückt sich auch in den Erwartungen von Lehrkräften aus. Mit dem Einsatz von Experimenten im Unterricht verfolgen Lehrkräfte sehr unterschiedliche Ziele. Als übergeordnete Ziele werden dabei häufig die Vermittlung von Fachwissen und die Steigerung des Interesses am Fach genannt (Welzel et al., 1998; Karaböcek & Erb, 2014).

Obwohl eine ganze Reihe an Studien zur Wirksamkeit des Experimentierens zur Verfügung stehen (vgl. den folgenden Abschnitt) sind die Effekte einzelner Methoden nicht zufriedenstellend aufgeklärt. Dies wird in der Regel damit in Verbindung gebracht, dass Elemente auf der Ebene der Sichtstruktur des Unterrichts nur wenig Einfluss auf die Wirksamkeit des Unterrichts im Ganzen haben. Der Grund hierfür ist zum einen, dass durch das Handeln der Lehrkräfte die für die Treatments in Vergleichsstudien vorgesehenen Unterrichtsskripte nicht immer genau wie vorgesehen im Unterricht umgesetzt werden. Zum anderen ist zu erwarten, dass Elemente der Tiefenstruktur, wie das Handeln der Lehrkräfte im weiteren Umfeld der Treatments Effekte hervorrufen, die mögliche Effekte verschiedener Unterrichtssituationen überlagern.

Demgegenüber aber sehen Konzepte für fachdidaktische Forschung, die auf eine Verbesserung des naturwissenschaftlichen Unterrichts gerichtet ist, die Notwendigkeit, auch Unterrichtsskripte zu entwerfen und zu untersuchen, die den Erfolg des Unterrichts erhöhen könnten. So erkennen Klieme & Rakoczy, dass sich fachdidaktische Empfehlungen „auf bewährte professionelle Unterrichtsprinzipien stützen“ können: „Empirische Legitimation können sie jedoch erst aus Längsschnittstudien erhalten, in denen der Einfluss der Unterrichtsgestaltung (...) auf den Kompetenzzuwachs der Schülerinnen und Schüler untersucht wird“ (Klieme & Rakoczy, 2008).

Dementsprechend sieht beispielsweise auch das Rahmenmodell für PISA 2015 einen Blick auf die relevanten Komponenten des naturwissenschaftlichen Unterrichts vor und nennt als ein Element der Instruktionsqualität domänenspezifische Aktivitäten, wie konkrete Beispiele aus dem Bereich „Science Teaching and Learning“ (Müller, Prenzel, Seidel, Schiepe-Tiska, & Kjærnsli, 2016).

Der vorliegende Beitrag stellt die Ergebnisse einer umfangreichen empirischen Studie vor, in der die Auswirkungen der Durchführung von Schülerexperimenten auf die Entwicklung des Fachwissens von Schülerinnen und Schülern der siebten Klasse mit den Auswirkungen einer Lernumgebung mit Demonstrationsexperimenten verglichen wurde. Um trotz der genannten Schwierigkeiten zu einer Aussage über die Wirksamkeit der hier in Frage stehenden Unterrichtsmethoden zu kommen, wurden zum einen Treatments entwickelt und realisiert, die den Einfluss des individuellen Handelns der Lehrkräfte möglichst geringhalten. Da die Treatments aber immer nur Teil eines umfangreicheren Unterrichts sind, der Einfluss auf den Erfolg der Arbeit der Schülerinnen und Schüler hat – und zwar auch bei zeitnaher Durchführung der Pre- und Posttests –, haben wir weiter die Lehrkräfte als Faktor in die Datenanalyse einbezogen, hier zunächst explorativ und vorbereitend für weitere Untersuchungen.

2. Forschungsstand

2.1. Experimentieren im Physikunterricht – eine Definition

Der Begriff des Experimentierens im Zusammenhang mit naturwissenschaftlichem Unterricht wird häufig verwendet, wobei aber durchaus verschiedene Vorstellungen darüber vorzufinden sind, welche Tätigkeiten hierunter zu verstehen sind.

In der naturwissenschaftlichen Forschung ist Experimentieren implizit oder explizit hypothesengesteuert; in der Regel ist es ergebnisoffen. Mit Experimenten im Unterricht wird dagegen kein Neuland für die Naturwissenschaften betreten, wohl aber für die beteiligten Schülerinnen und Schüler. Während für diese Tätigkeit oft der Begriff „Versuch“ vorgeschlagen wird, sehen andere Autorinnen und Autoren den Begriff „Experiment“ als gerechtfertigt an, wenn eine Fragestellung explizit mit dieser Tätigkeit verbunden ist. Die Vielzahl möglicher Experimentalstrategien im Unterricht, bei der der forschend-entdeckende Unterricht eine zentrale Rolle einnimmt, beschreiben Höttecke & Rieß (2015) im Rahmen einer Betrachtung aktueller Wissenschaftsforschung über das Experimentieren in Wissenschaft und Schule. Eine häufig verwendete Systematisierung stellt das explorative dem theoriegeleiteten Experimentieren gegenüber. Die Fokussierung auf die Minimierung von Messfehlern sowie deren Kontrolle stellt eine weitere Strategie dar. Auch experimentelle Variationen zur eingehenderen Untersuchung eines Phänomens sowie die grundsätzliche

Wiederholung zum Zweck der Bestätigung werden in diesem Zusammenhang genannt. Höttecke & Rieß (2015) verweisen explizit darauf, dass es nur genau eine Definition für das Experiment nicht geben kann.

Bezugnehmend auf Hacking (1989) wird für die vorliegende Arbeit die Tätigkeit des Experimentierens im Unterricht definiert als “an activity which involves an intervention to produce the phenomenon to be observed or to test a hypothesis” (Hacking, 1989 in: Millar, 2010, 109). Reines Beobachten, so dass die Aktivität zur planvollen Erzeugung eines Phänomens fehlt – zum Beispiel bei der Beobachtung eines Naturphänomens – ist nach dieser Definition kein Experimentieren. Gleichmaßen gilt für die vorliegende Studie, dass das Ziel des Experimentierens nicht darin bestehen soll, Gegenstände beispielsweise aus rein spielerischen Gesichtspunkten geplant zu variieren. Vielmehr soll das Ziel des Experimentierens ein verbessertes Verständnis des zugrundeliegenden Sachverhalts sein, was sich in einer messbaren Veränderung des Fachwissens zeigt.

Dies stellt ein engeres Verständnis von Experimentieren dar als von Lunetta, Hofstein & Clough (2007) vorgeschlagen. Im Gegensatz zu dem dort formulierten Verständnis wird in der vorliegenden Arbeit zum Beispiel das Auswerten experimenteller Daten, die von der Lehrkraft zur Verfügung gestellt werden, nicht als Experimentieren verstanden. Gleichzeitig bietet die oben genannte Definition Raum für Experimentiersituationen, die sich leicht in der Schulpraxis wiederfinden lassen.

Darüber hinaus lassen sich für den Einsatz von Experimenten noch eine ganze Reihe anderer Ziele finden (Karaböcek & Erb, 2015), wie etwa eine höhere Authentizität des Physikunterrichts durch Repräsentation von Elementen, die zum Methodenrepertoire der Physik gehören (Höttecke & Rieß, 2015).

2.2. Erwartungen an das Experimentieren im naturwissenschaftlichen Unterricht

Wie Welzel et al. (1998) im Rahmen einer europäischen Delphi-Studie zeigen konnten, sind viele Lehrkräfte zuversichtlich, dass ihre Schülerinnen und Schüler insbesondere durch selbstständiges Experimentieren in Kleingruppen ein tieferes Verständnis und damit eine Verbindung zwischen Theorie und Praxis – wie zum Beispiel von Hodson (1993) gefordert – erreichen. Im Rahmen einer Befragung von britischen Naturwissenschaftslehrkräften wird von diesen die Bedeutung des Experimentierens herausgehoben (National Endowment for Science, Technology and the Arts [NESTA], 2008). Lynch und Ndyetabura (1993) befragten 257 Lehrkräfte zu ihren Zielen, wenn sie in ihrem Unterricht „practical work“ einsetzen: Für die meisten von ihnen diene das Experimentieren der Verbesserung eines theoretischen Verständnisses. Auch Schülerin-

nen und Schüler selbst äußern ähnliche Erwartungen an das Schülerexperiment (Behrendt, 1991; Winkelmann & Erb, 2011).

Eine Lehrkräftebefragung über alle Unterrichtsfächer hinweg zur Anwendungshäufigkeit verschiedener Unterrichtsmethoden kommt zu dem Ergebnis, dass zwar keine „unterrichtsmethodische Monstruktur“ nachgewiesen werden könne, allerdings „bei gleichzeitiger Dominanz des Frontalunterrichts [...] eine unterrichtsmethodische Vielfalt auf geringem Häufigkeitsniveau erkennbar“ sei (Bohl, 2001, 279). Für den naturwissenschaftlichen Unterricht kommt Tesch (2005) mit einer Videostudie zur Unterrichtsrealität im deutschen Physikunterricht zu etwas anderen Erkenntnissen. Tesch konnte zeigen, dass etwa zwei Drittel der Unterrichtszeit für das Experimentieren, inklusive Vor- und Nachbereitung, verwendet werden. Dabei dominiert als Unterrichtsmethode das Schülerexperiment mit zwei Dritteln der reinen Experimentierzeit – wobei häufig nicht von selbstständigem Experimentieren, sondern von einem Abarbeiten einer detaillierten Anleitung ausgegangen werden muss. Im Hinblick auf die darauf verwendete Unterrichtszeit ist damit belegt, dass das Experimentieren eine bedeutende Rolle im Physikunterricht einnimmt. Tesch bietet in ihrer Studie ein Kategoriensystem zur Beschreibung von Experimentiersituationen im Physikunterricht an. Darauf aufbauend vergleicht Börlin (2012) Physikunterricht in Finnland, der Schweiz und in Deutschland. Beide verwenden in ihrem Monitoring Kategorien, die in der vorliegenden Studie für die Definition der geplanten Experimentiersituationen genutzt wurden.

Die bisher genannten Studien gehen davon aus, dass das Experimentieren erfolgreich ist und man einen Effekt – etwa Leistungszuwachs oder Veränderungen bei Interesse und Motivation – durch das Experimentieren erwarten kann. Einen etwas anderen Blick auf das Experimentieren legen Schecker, Neumann, Theyßen, Eickhorst & Dickmann (2016). Sie gehen der Frage nach, welche Fähigkeiten Schülerinnen und Schüler im Hinblick auf das Experimentieren besitzen und versuchen so das Konstrukt einer experimentellen Kompetenz zu definieren, um diese dann messen zu können (siehe auch: Theyßen, Schecker, Neumann, Eickhorst, & Dickmann, 2016).

Einen breiten Forschungsüberblick bezüglich allgemein forschenden Handelns im naturwissenschaftlichen Unterricht bieten Rönnebeck, Bernholt & Rohpohl (2016). Unter forschendem Handeln („scientific inquiry activities“) verstehen sie dabei neun verschiedene Bereiche, in die sich das Experimentieren einordnen lässt. Unsere Studie lässt sich anhand dieser Kategorisierung den Bereichen „3. Hypothesen formulieren und Vorhersagen treffen“ und „4. Planung und Durchführung von Untersuchungen zuordnen“.

2.3. Schülerexperimente und Demonstrationsexperimente im naturwissenschaftlichen Unterricht

Betrachtet man Studien zum Experimentieren im schulischen Umfeld, so ergibt sich ein nicht eindeutiges Ergebnis zum Einfluss von Schülerexperimenten oder Demonstrationsexperimenten. Die naturwissenschaftsdidaktische Forschung beschäftigt sich verstärkt seit den 1960er Jahren mit dem Thema des selbstständigen Experimentierens. Auffällig ist, dass die Mehrzahl der Studien zum Einfluss des Experimentierens den Lernfortschritt in den Fokus rückt und nicht die Entwicklung von Interesse, Motivation und Haltungen zur Naturwissenschaft (Potvin & Hasni, 2014; White, 1996).

Natürlich kann praktisches naturwissenschaftliches Arbeiten in der Schule ganz unterschiedliche Wirkungen haben, wie Hofstein und Mamlok-Naaman (2007) berichten: „Inquiry-type laboratories have the potential to develop students’ abilities and skills such as: posing scientifically oriented questions [...], forming hypothesis, designing and conducting scientific investigations, formulating and revising scientific explanations, and communicating and defending scientific arguments” (S. 105, siehe auch Blanchard et al., 2010).

Umfangreiche Reviews zum Thema „Lernen in den Naturwissenschaften“ zeigen, dass „laboratory work“ lediglich kleine Beiträge zu einem verbesserten Verständnis von Naturwissenschaften zu leisten vermag (Shulman & Tamir, 1973; Hofstein & Lunetta, 1982; Singer, Hilton, & Schweingruber, 2006). Diese Überblicksbeiträge erfassen den Einfluss des praktischen Arbeitens auf das Gesamtverständnis von Naturwissenschaften, also zum Beispiel auch Wissen über die Arbeitsweisen dieser Wissenschaften. Ein aktuelleres Review mit stärkerem Fokus auf die Wissensvermittlung naturwissenschaftlicher Inhalte findet sich dagegen bei Millar (2010). Andere Studien, bei denen der Einsatz von Schüler- bzw. Demonstrationsexperimenten direkt verglichen wurde (z. B. Hopf, 2007) lassen erwarten, dass Unterschiede in der Unterrichtsorganisation, also auf der Ebene der Sichtstruktur, nicht entscheidend zumindest für den Unterrichtserfolg in fachlicher Hinsicht sind. Entscheidender sind vielmehr tieferliegende Prozesse und damit verbunden die Lehrperson selbst (z. B. auch Hattie, 2013).

Zum einen lassen sich viele Studien zum Thema nicht miteinander vergleichen. Dies liegt unter anderem an

- unterschiedlichem Sprachgebrauch darüber, wie Experimente im Unterricht eingesetzt werden (inquiry learning, laboratory work, practical work, hands-on-Situationen, Experimentiersituationen, ...),
- unterschiedlichen Inhaltsfeldern der Naturwissenschaften (Abrahams & Miller, 2008; Chemie: Anders et al., 2003) sowie den einzelnen Fach-

domänen (Mechanik: Thijs und Bosch, 1995), Gasgesetze (Atkinson, 1981), E-Lehre und Optik (Hopf, 2007),

- unterschiedlichen unabhängigen Variablen: z. B. Grad des Schülerhandelns (z. B. vorliegende Studie), Art der Repräsentation (Schwichow, Zimmermann, Croker & Härtig 2016), Bedeutung der Nachbearbeitung von Experimenten (Muth & Erb, 2017), Problemorientierung (Hopf, 2007), Nutzung von Feedback-Systemen (Schulz, 2011),
- unterschiedlichen abhängigen Variablen: z. B. Fachwissen (Hofstein & Lunetta, 1998; Rumann, 2005; Blanchard et al., 2010) und Interesse (Schulz, 2011; Potvin, Hasni, & Sy, 2017) sowie
- nur schwer vergleichbaren Forschungsdesigns: z. T. unklarer Bericht des Designs, sehr kleine Stichproben, mangelnde Kontrolle von Störvariablen, nur schwer vergleichbare Landesspezifika (Watson & Prieto & Dillon, 1995).

Zum anderen wird in der Literatur insbesondere darauf hingewiesen, dass sich die Ergebnisse vieler Studien oft widersprechen oder zum Teil gar wissenschaftsmethodische Mängel aufwiesen (Garrett & Roberts, 1982; Thompson & Soyibo, 2002; Hofstein & Lunetta, 2004). Für uns ist genau dies der Anlass, die hier beschriebene Studie unter möglichst kontrollierten Bedingungen und Darlegung aller Randbedingungen durchzuführen.

Potvin und Hasni (2014) beschreiben die Lage in ihrem Review ähnlich und betonen zusätzlich, dass überhaupt nur wenige Studien „practical work“ als reinen Untersuchungsgegenstand haben. Das macht es schwierig, zum Teil positiv berichtete Effekte von etwaigen anderen Bestandteilen der Untersuchungen zu trennen.

2.4 (Aktuelles) Interesse an Physik

Für den deutschsprachigen Raum berichten Hoffmann et al. (1998), dass gute schulische Leistungen mit einem hohen Interesse an Physik einhergehen. Auch international stellen Potvin und Hasni (2014) in ihrem Review über Studien zu Interesse, Motivation und Einstellungen in Naturwissenschaften- und Technikfächern fest, dass die meisten gefundenen Studien von einem positiven Zusammenhang zwischen Interesse, Motivation und Einstellungen auf der einen Seite und der schulischen Leistung auf der anderen Seite berichten.

Das „Interesse an Physik“ gilt als recht stabile Eigenschaft und sollte eher über lange Zeiträume, z. B. die komplette Schullaufbahn hinweg, beobachtet werden. In unserer Studie wird als weitere abhängige Variable das aktuelle Interesse von Schülerinnen und Schülern an Physik gemessen. Nach Schulz (2011) handelt es sich hierbei um situationsbedingtes Interesse an z. B. einem Unterrichtsprozess.

3. Forschungsfragen und Hypothesen

Die Forschungsfrage sowie die Anlage der vorliegenden Studie reagieren auf den beschriebenen Klärungsbedarf. Die Absicht der Studie war es, einerseits herauszufinden, welche Auswirkungen das Experimentieren auf die Leistung von Schülerinnen und Schülern hat. Zusätzlich wurde als affektives Merkmal das Interesse an Physik erhoben, da auch hier Zusammenhänge vermutet werden dürfen. Ein besonderes Augenmerk liegt hierbei auf dem aktuellen Interesse an Physik.

Für den Lernfortschritt (AV) der Schülerinnen und Schüler gingen wir davon aus, dass dieser größer ist, wenn die Schülerinnen und Schüler selbstständig experimentieren anstatt der Demonstration durch die Lehrkraft folgen.

Ebenso erwarteten wir, dass das aktuelle Interesse an Physik (AV) ausgeprägter ist, wenn die Schülerinnen und Schüler selbstständig experimentieren.

4. Methode der Untersuchung

4.1. Design

Bei der hier beschriebenen Untersuchung handelt es sich um eine quasi-experimentelle Vergleichsstudie mit drei Treatments im Physikunterricht der Sekundarstufe I. Verglichen wurden drei Methoden der Einbindung naturwissenschaftlichen Experimentierens, bei denen die Schüleraktivität (Faktor) variiert.

Die Schüler experimentieren in kleinen Gruppen oder schauen ihrer Lehrkraft oder einem einzelnen Schüler beim Experimentieren zu. Da die Treatments in den normalen Unterrichtsprozess eingebunden wurden, konnten die Schülerinnen und Schüler nicht randomisiert auf sie verteilt werden, sondern verblieben in ihren jeweiligen Klassenverbänden – es wurden also jeweils vollständige Klassen einem Treatment zugeordnet. Die Zuordnung der Klassen zu den Treatments erfolgte zufällig. Die Intervention umfasste zwölf Unterrichtsstunden zur geometrischen Optik und hatte damit eine Dauer von sechs Wochen. Da der Einfluss der Experimentiersituationen untersucht werden sollte, wurde auf Übungs- und Wiederholungsphasen verzichtet, um vergleichbare Rahmenbedingungen zu gewährleisten. Der dabei typischerweise realisierte Unterricht begann mit einer Einführung in die Fragestellung (5 Min.), dann begann die Experimentierphase (ca. 30 Min.). Die Schülerinnen und Schüler erhielten hierbei die nötigen Materialien, das weitere Handeln wurde durch Arbeitsblätter vorstrukturiert. Den Abschluss bildete das gemeinsame Klassengespräch zur Sicherung der im Experiment gemachten Beobachtungen (ca. 10 Min.).

Inhaltlich wurden über mehrere Stunden verschiedene, schulübliche Experimente zur geometrischen Optik bearbeitet. Tabelle 1 gibt einen Überblick über den Verlauf der Unterrichtsreihe. Eine eingehendere Darstellung des Unterrichtsverlaufs findet sich bei Winkelmann & Erb (2014).

Die Experimentieranleitungen sind verfügbar unter http://www.mnu.de/attachments/article/635/MNU_Heft_14_07_Winkelmann.pdf

	Thema	Experiment
1. Stunde	Untersuchungen von Spiegelbildern	1. Schatten vor und im Spiegel 2. Entfernung des Spiegelbildes im Spiegel
2. Stunde	Reflexionsgesetz	3. Licht, Lot und reflektiertes Licht 4. Betrachtung des Spiegelbildes zweier Personen
3. Stunde	Untersuchungen zur Lichtbrechung	5. Peil-Experiment: Fische jagen 6. Peil-Experiment: Stecknadeln in einer Reihe
4. Stunde	Brechungsgesetz und Totalreflexion	7. Lichtbrechung mit Tischoptik 8. Totalreflexion mit Tischoptik
5. Stunde	Licht dringt durch Sammellinse, Bedingungen für Lichtbündelung	9. Lichtbündelung mit Schusterkugel 10. Lichtbrechung an Sammellinse
6. Stunde	Bildentstehung durch Sammellinse	11. Experimentieren an der optischen Bank 12. Experimentieren an der optischen Bank

Tab. 1: Überblick der verwendeten Experimente

Einige der Experimente haben qualitativen Charakter, andere Experimente haben ihren Fokus auf dem quantitativen Messen von Winkeln. Alle Experimente wurden als Realexperimente in kleinen Schülergruppen (für gewöhnlich waren drei Schülerinnen bzw. Schüler in einer Gruppe) oder durch die Lehrkraft als Demonstration durchgeführt. In einigen Experimenten wurden für die Schülerinnen und Schüler neue physikalische Beobachtungen eingebracht. Andere Experimente dienen der Vertiefung des vorhandenen Wissens. Nimmt man Kategorien etwas älterer internationaler Studien zu Hilfe (Etkina, Van Heuvelen, Brookes & Mills, 2002), fallen die meisten unserer Experimente in den "Typus 1": die Schülerinnen und Schüler machen Beobachtungen und analysieren diese im Nachgang des Experiments. In einem der Experimente mussten die Schülerinnen und Schüler darüber hinaus explizit ihre eigene Hypothese mit Hilfe des Experiments testen. Eine ausführlichere Darstellung der verwendeten Experimente kann bei Winkelmann (2015) nachgelesen werden.

Wie weiter oben ausgeführt, wird vermutet, dass das Experimentieren auch einen Einfluss auf affektive Merkmalsausprägungen wie das Interesse an Physik haben wird. Daher werden mit der geplanten Intervention auch andere als rein fachwissenschaftliche Funktionen untersucht, jedoch sind die verwendeten Experimente nicht explizit darauf ausgerichtet.

4.2. Stichprobe

Insgesamt nahmen an der Studie 1032 Schülerinnen und Schüler der siebten Jahrgangsstufe aus Deutschland teil. Allerdings konnten sich nicht immer alle Schülerinnen und Schüler an jedem Messzeitpunkt beteiligen, etwa wegen einer Krankheit. Wenn aus solchen Gründen Schülerinnen und Schüler zwei oder mehr Unterrichtsstunden verpassten, wurden deren Daten nicht in der Auswertung berücksichtigt. Das stellt wegen der zufälligen Verteilung kein Problem dar. Abschließend bildeten 858 Schülerfragebögen die Grundlage für die Auswertung.

Die 408 Schüler und 425 Schülerinnen (25 gaben ihr Geschlecht nicht an) waren im Schnitt 12,2 Jahre alt ($SD = 0.54$) und besuchten alle ein Gymnasium. Die Schülerinnen und Schüler waren auf 39 Klassen und 23 Lehrkräfte verteilt. Einige Lehrkräfte haben Klassen parallel unterrichtet. Eine Übersicht der Verteilung bietet Tabelle 2. Die teilnehmenden Lehrkräfte haben dabei selbst entschieden, in welcher Experimentiersituation sie unterrichten wollen. Dadurch wurde sichergestellt, dass Lehrkräfte sich nicht mit der ihnen zugedachten Experimentiersituation überfordert fühlen.

Lehrkraft	Experimentiersituation		
	A	B	C
1	X	X	X
2		X	
3	X	X	
4	X		X
5	X		X
6		X	X
7	X		X
8	X		X
9	X	X	
10		X	X
11		X	
12	X		X
13		X	X
14	X		X
15	X	X	
16		X	
17	X		
18			X
19		X	
20		X	X
21			X
22	X		
23		X	X

Tab. 2: Verteilung der Lehrkräfte auf die Experimentiersituationen

Um die Einhaltung der jeweiligen Experimentiersituation im Unterricht sicherzustellen, erhielten die Lehrkräfte zusätzlich zu den strukturgebenden Arbeitsblättern Unterrichtsskripte. Die Lehrkräfte wurden vor der Durchführung der Studie eigens geschult, damit die Umsetzung der Treatments zuverlässig erfolgt, d. h. jede Lehrkraft die ihr zugeteilte Experimentiersituation skriptgetreu realisiert.

Wenn zu vermuten ist, dass eine weitere Einflussgröße mit der eigentlich interessierenden unabhängigen Variable konfundiert bzw. sich zwei Effekte überlagern, lässt sich dies in einem faktoriellen Forschungsplan mit Hilfe einer zusätzlichen unabhängigen Variable, der sogenannten „Confoundervariable“ kontrollieren. Diese Variable wird dann als zweiter Faktor in den Versuchsplan mit aufgenommen (Bortz & Döring, 2006). Für die vorliegende Arbeit wurde als kontrollierender Faktor der Einfluss der unterrichtenden Lehrkraft berücksichtigt (als zufälliger Faktor in der Varianzanalyse ANOVA).

4.3. Die unterschiedlichen Experimentiersituationen

Die Untersuchung vergleicht drei verschiedene Experimentiersituationen auf ihren Einfluss auf das Fachwissen und das aktuelle Interesse an Physik. Um den Einfluss der Ausgestaltung der Anleitungen gering zu halten, wurden diese inhaltlich hochgradig ähnlich, in dem Grad der Detailliertheit jedoch unterschiedlich gestaltet. Die drei Situationen (a-c) sind die unabhängigen Variablen und konzeptionell auf Grundlage der oben beschriebenen Literatur begründet:

- a) Die Lehrkraft oder einzelne Schülerinnen oder Schüler demonstrieren ein Experiment vor der gesamten Klasse.
- b) Die Schülerinnen und Schüler führen eigenständig Experimente in Kleingruppen durch und erhalten hierfür detaillierte Instruktionen. Diese Anleitungen werden in geschriebener Form den Schülerinnen und Schülern zur Verfügung gestellt.
- c) Die Schülerinnen und Schüler führen eigenständig Experimente in Kleingruppen durch und erhalten hierfür offene Instruktionen. Diese sind nicht auf ein Minimum beschränkt, aber weniger ausgearbeitet als in Gruppe B. Eine Problemsituation wird benannt und die zur Verfügung stehenden Materialien aufgelistet. Der Aufbau des Experiments und ggf. das Formulieren von Hypothesen muss selbstständig geleistet werden.

Die verschiedenen Phasen des experimentellen Arbeitens werden in Tabelle 3 zusammengefasst und gründen auf anderen Studien in den Vereinigten Staaten (z. B. Klahr, Zimmermann & Jirout, 2011) und Deutschland (z. B. Schreiber, Theyßen & Schecker, 2009).

Phasen des Experimentierens	
Planung	Problemstellung Entwicklung des Experiments Formulierung einer Hypothese
Durchführung	Versuchsanordnung aufbauen Messen und Handhabung von Material Dokumentieren
Auswertung	Messdaten verarbeiten Interpretation der Daten

Tab. 3: Phasen des Experimentierens

In der vorliegenden Studie fanden die Variationen in den Phasen Planung und Durchführung eines Experiments statt. Um einen Einblick in die einzelnen Phasen zu geben, sind in der rechten Spalte einige Tätigkeiten benannt.

Abbildung 1 fasst die Experimentiersituationen zusammen. Zusätzlich wird in den Stufen Planung und Durchführung genau gezeigt, an welcher Stelle die studienbedingten Variationen stattfanden.

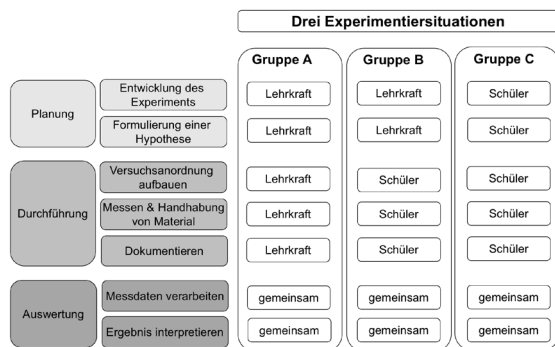


Abb. 1: Die drei Experimentiersituationen unterscheiden sich in der Schüleraktivität während der Planung und Durchführung des Experimentierens.

Die ebenfalls bedeutsame Phase der Auswertung wurde stets von den Lehrkräften und Schülerinnen und Schülern gemeinsam durchlaufen. Zur Auswertungsphase liegt eine aktuelle umfassende Studie statt (Muth, 2018).

4.4. Variablen und Messinstrumente

4.4.1. Physikalisches Fachwissen

Die abhängige Variable zur ersten Forschungsfrage ist das physikalische Fachwissen. Für deren Erfassung wurde ein Test konstruiert, dessen Items sich auf die fachlichen Inhalte der Experimentiersituationen beziehen. Insgesamt wurden 34 Aufgaben entwickelt und in Winkelmann (2015) veröffentlicht. Nicht alle Aufgaben konnten für die Analyse des Fachwissens im Rahmen der Studie genutzt werden.

Die in den Experimentiersituationen angesprochenen Inhalte und damit die im Fachwissenstest verwendeten Aufgaben orientieren sich an den Anforderungen üblicher Lehr- bzw. Bildungspläne, wodurch eine curriculare Validität gegeben ist.

derungen üblicher Lehr- bzw. Bildungspläne, wodurch eine curriculare Validität gegeben ist.

Zusätzlich wurden im Rahmen einer wissenschaftlichen Hausarbeit Fallanalysen durch Interviews mit teilnehmenden Schülerinnen und Schülern durchgeführt (OpIt, 2013). Dabei wurde der Frage nachgegangen, inwieweit für die Schülerinnen und Schüler der Aufgabentext verständlich ist und ob sie den Arbeitsauftrag so interpretieren, wie er von den Aufgabenkonstruktoren vorgesehen ist. Die Ergebnisse lieferten wichtige Hinweise darauf, welche Aufgaben zur Analyse des Fachwissens der Schülerinnen und Schüler herangezogen werden konnten.

Üblicherweise wird für die Erhebung von Wissen auf Multiple-Choice-Tests zurückgegriffen (Haladyna, 1994). Durch die Wahl des geschlossenen Antwortformats „single-select“ mit drei Distraktoren und einem Attraktor wird eine hohe Objektivität in der Beurteilung der Schülerantworten sowie eine geringe Ratewahrscheinlichkeit von 25% erreicht. Für den verwendeten Fachwissenstest kann Cronbachs α als Gütemaß für die Reliabilität verwendet werden. Allerdings ist das zugrunde gelegte Konstrukt „geometrische Optik“ sehr weit gefasst, weswegen die Aussagekraft von Cronbachs α als eingeschränkt erachtet werden muss. Für den von uns verwendeten Test wurde ein Cronbachs α von .602 berechnet, was im Falle von Gruppenvergleichen als ausreichend erachtet wird (Lienert & Raatz, 1998).

Um eine Vorstellung von der Art der Aufgaben zu vermitteln, werden in Tabelle 4 zwei Beispielaufgaben gezeigt. Die erste dieser Aufgaben floss nicht in die Analyse mit ein, da im Einleitungstext die Bedingung des schräg auf die Grenzfläche treffenden Lichts nicht genannt wurde. Die zweite Aufgabe konnte verwendet werden; im vorangegangenen Unterricht wurde von den Schülerinnen und Schülern untersucht, was mit Licht geschieht, wenn es durch eine wassergefüllte Kugel leuchtet. Im Anschluss wurde hypothesengestützt die Abhängigkeit von der Form des durchleuchteten Körpers untersucht.

Was kannst du über einen Lichtstrahl beim Übergang von Luft in einen anderen durchsichtigen Stoff (z. B. Glas oder Wasser) sagen?

- Der Lichtstrahl reicht bis in die Mitte des anderen Stoffes und bringt diesen zum Leuchten.
- Der Lichtstrahl ändert seine Richtung, sobald er den anderen Stoff erreicht hat.
- Der Lichtstrahl geht nur bis zum Übergang zum anderen Stoff und hört dann auf.
- Der Lichtstrahl verändert seine Richtung nicht, sondern verläuft einfach gerade weiter.

Wie muss ein Gegenstand geformt sein, damit er Licht bündelt?

- Der Gegenstand muss rund wie eine vollständige Kugel geformt sein.
- Der Gegenstand muss am Rand dünner sein als in der Mitte.
- Der Gegenstand muss am Rand dicker sein als in der Mitte.
- Der Gegenstand muss überall gleich dick sein.

Tab. 4: Zwei Beispielaufgaben aus dem Fachwissenstest, wobei die erste Aufgabe nicht in die Auswertung mit einfließt.

4.4.2. Aktuelles Interesse an Physik

Als weitere abhängige Variable wurde direkt im Anschluss an die Unterrichtsstunden das aktuelle Interesse an Physik erhoben. Die Skala zum aktuellen Interesse an Physik geht auf Hoffmann et al. (1998) zurück, wurde in der Chemiedidaktik von Schulz (2011) aufgegriffen und für die vorliegende Arbeit für das Fach Physik angepasst. Insgesamt umfasst das Konstrukt neun Items, die von den Schülerinnen und Schülern auf einer fünfstufigen Ratingskala bewertet werden sollten. Das aktuelle Interesse an Physik bezieht sich auf den zuletzt erlebten Physikunterricht und wurde im Rahmen des Posttests erhoben. Die Reliabilität lag bei Cronbachs $\alpha = .86$. Ein Beispielitem lautet: „In den vergangenen Stunden ist die Zeit sehr schnell vergangen“.

4.4.3. Kontrollvariablen

Um eine Aussage über die Homogenität der Zusammensetzung der Vergleichsgruppen treffen zu können, wurden auf fachlicher Ebene eine Skala (Q2) zu kognitiven Fähigkeiten (Heller & Perleth, 2000) sowie die letzten Zeugnisnoten der Schülerinnen und Schüler in den Fächern Physik, Mathematik und Deutsch erfasst. Weiter kam eine Skala zum physikalischen Selbstkonzept (Hoffmann et al., 1998) zum Einsatz.

Im Pretest wurde die allgemeine kognitive Fähigkeit der Schülerinnen und Schüler durch die Subskala Q2 – einem quantitativen Teil – des Kognitionsfähigkeits-Tests (KFT, Testheft 2) von Heller und Perleth (2000) gemessen. Hierbei handelt es sich um sogenannte Powertests, bei denen die Schülerinnen und Schüler gebeten werden, in einer bestimmten Zeit so viele Aufgaben wie möglich zu bearbeiten.

Da dieser Test eine gute Langzeitstabilität aufweist, genügt eine einmalige Erhebung zu Beginn der Studie.

4.5. Erhebung

Die Studie wurde in einem Pre-/Posttest-Design durchgeführt. Der Fokus der Studie lag auf möglichen Unterschieden in der Entwicklung des physikalischen Verständnisses innerhalb der drei oben beschriebenen Experimentalsituationen. Hierfür wurde eine Unterrichtsreihe entworfen und diese den teilnehmenden Lehrkräften vorab vorgestellt.

Abbildung 2 fasst den Ablauf der Studie sowie die drei Situationen des Experimentierens zusammen.

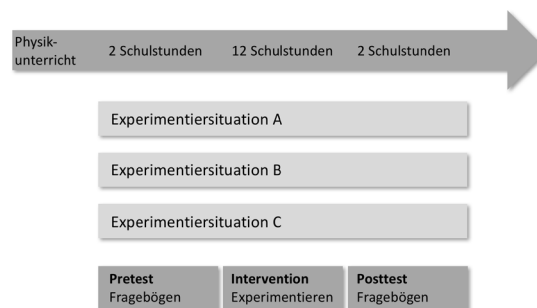


Abb. 2: Erhebungsablauf mit den drei Vergleichsgruppen A, B und C

4.6. Auswertung mittels Varianzanalyse

Die Datenauswertung erfolgte mit Hilfe einer Varianzanalyse, um die drei Ausprägungen des Treatmentfaktors miteinander vergleichen zu können. Ein weiterer Grund für die Nutzung von Varianzanalysen war die Möglichkeit, den Einfluss des zufälligen Faktors „Lehrkraft“ zu erfassen. Es handelt sich daher um eine zwei-faktorielle Varianzanalyse (ANOVA) im mixed-design.

Um (insbesondere bei kleinen Stichproben) eine ANOVA sinnvoll durchführen zu können, ist es notwendig, dass verschiedene Voraussetzungen erfüllt sind. Diese sind in der Regel eine Prüfung des Vorliegens einer Normalverteilung der auszuwertenden Daten mit Angabe der Schiefe und Kurtosis der Verteilung sowie eine Prüfung auf Varianzhomogenität.

Für eine ANOVA gelten zusätzlich zwei weitere Annahmen: Die Messungen müssen unabhängig voneinander sein, und die abhängige Variable sollte mindestens intervallskaliert sein (Field, 2011, S. 359). Beide zusätzlichen Annahmen sind in der vorliegenden Studie erfüllt.

Üblicherweise werden Unterschiedshypothesen durch den Vergleich mindestens zweier Gruppen getestet. In diesem einfachsten Fall würde eine behandelte Treatment-Gruppe mit einer unbehandelten Kontroll-Gruppe verglichen (Bortz & Döring, 2006). In der hier beschriebenen Studie bedeutet der Begriff „behandelte Gruppe“, dass auf eine bestimmte Art und Weise im Unterricht experimentiert wird. Der

Kontrollgruppe würde man jegliche Experimente vorenthalten. Da dies jedoch in der kontrollierten Situation zu einer gravierenden Verzerrung der Unterrichtsrealität führen würde, wurde auf die Bildung einer Kontrollgruppe verzichtet. Stattdessen fiel die Entscheidung auf einen Mehrgruppenplan, bei dem drei Treatments unterschiedlicher Ausprägung miteinander verglichen werden. Diese Treatments bildeten als Faktorstufen gemeinsam als unabhängige Variable den festen Faktor „Experimentiersituation“ für die Varianzanalyse.

Die interne Validität einer quasi-experimentellen Untersuchung wie der vorliegenden ist gefährdet, wenn sich die Probanden, hier die Schülerinnen und Schüler, nicht nur in Bezug auf die unabhängige Variable (Experimentiersituationen A-C), sondern auch in anderen personenbezogenen (abhängigen) Variablen unterscheiden (Bortz & Döring, 2006; Tabachnick & Fidell, 2007). Als eine abhängige Variable ist in dieser Studie das Fachwissen der Schülerinnen und Schüler von Interesse. Es zeigte sich, dass die Gruppen in den einzelnen Treatments bereits im Pretest unterschiedliche Leistungen aufwiesen. Um diesen Unterschieden im Pretest Rechnung zu tragen, wird im Rahmen der Varianzanalyse als abhängige Variable die Differenz zwischen dem Posttest und dem Pretest verwendet (Bortz & Döring, 2006).

Alternativ ist auch eine ANOVA mit Messwiederholung denkbar. Allerdings hat diese den Nachteil, dass hierbei der zu vermutende Einfluss der unterrichtenden Lehrkräfte nicht als zufälliger Faktor mitberücksichtigt werden kann. Zur Kontrolle wurden dennoch stets auch solche Varianzanalysen durchgeführt.

Eine andere Möglichkeit, die ungleichen Ausgangswerte zu berücksichtigen, stellt die Durchführung einer Kovarianzanalyse dar. Da sich die Pretestleistungen der einzelnen Treatments jedoch signifikant voneinander unterscheiden, darf eine solche Analyse mit den Leistungen im Pretest als Kovariate nicht gerechnet werden (Field, 2011). Eine weitere Möglichkeit, stellen sogenannte „matched samples“ dar, bei denen die Gruppengrößen gleich groß sind. Erzeugt man das „matched sample“ auf Grundlage der Leistungen im Pretest, wird damit sichergestellt, dass in allen Treatments gleich viele Schülerinnen und Schüler mit gleicher Leistung im Pretest vertreten sind. Unterschiedliche Leistungen zum Zeitpunkt des Pretests werden in dieser Arbeit grundsätzlich mit der Bildung der Lernzuwächse bzw. der Entwicklung der affektiven Schülermerkmale berücksichtigt. Das „matched sample“ stellt damit eine zusätzliche Kontrolle dar, die in dieser Studie dann zur Anwendung kommt, wenn statistische Robustheit verlangt ist – insbesondere bei Verletzung der Varianzhomogenitätsannahme.

5. Ergebnisse

5.1. Kontrollvariablen

Bezüglich der kognitiven Leistungen (KFT) zeigt eine ein-faktorielle ANOVA keine signifikanten Unterschiede zwischen den Vergleichsgruppen, $F(2,857) = .692$, $p = n. s.$ Das physikalische Selbstkonzept ist in allen drei Gruppen ähnlich ausgebildet. Die Mittelwerte variieren nicht signifikant, $F(2,854) = .470$, $p = n. s.$ Auch die letzten Zeugnisnoten verweisen auf eine hohe Homogenität innerhalb der drei Gruppen.

Bezogen auf die beschriebenen Kontrollvariablen kann daher davon ausgegangen werden, dass die an dieser Studie teilgenommenen Vergleichsgruppen mit ähnlichen Voraussetzungen gestartet sind.

5.2. Entwicklung des physikalischen Wissens

Der t-Test für abhängige Stichproben liefert für den Lernzuwachs aller Schülerinnen und Schüler – also die Entwicklung des Fachwissens zwischen Pre- und Posttest – ein höchst signifikantes Ergebnis, $t(858) = -18.763$, $p < .001$. Mit $\eta^2 = .59$ liegt die Effektstärke deutlich im Bereich eines großen Effekts.

Für die Schülerinnen und Schüler handelte es sich in der Regel um fachlich neue Inhalte der Physik. Daher ist es nicht erstaunlich, dass im Pretest im Mittel lediglich 4,7 Punkte von möglichen 17 Punkten erreicht wurden. Aus testtheoretischer Sicht ist das Abschneiden der Schülerinnen und Schüler während des Posttests im mittleren Punktebereich (8,6 Punkte) als gut zu erachten. Für die Varianzanalyse der Unterschiede zwischen den Vergleichsgruppen wurde eine zwei-faktorielle ANOVA ohne Messwiederholung mit dem festen Faktor „Experimentiersituation“ und dem zufälligen Faktor „Lehrkraft“ durchgeführt. Abbildung 3 zeigt den mittleren Lernzuwachs der drei Vergleichsgruppen. Die Mittelwerte der drei Vergleichsgruppen sind sehr ähnlich.

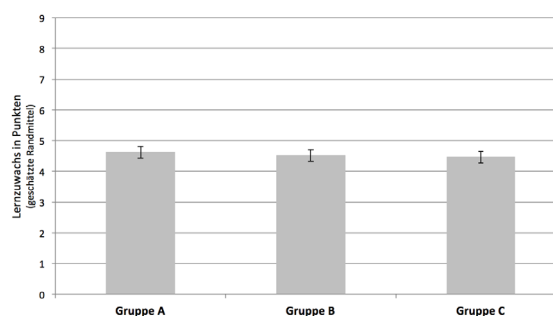


Abb. 3: Lernzuwachs aller Schülerinnen und Schüler in den verschiedenen Experimentiersituationen unter Berücksichtigung des zufälligen Faktors „Lehrkraft“, zusätzlich angegeben ist der Standardfehler.

Dementsprechend zeigt die ANOVA für den festen Faktor „Experimentiersituation“ auch keinen signifikanten Effekt, $F(2,821) = .093$, $p = n. s.$ Die Kontrollanalyse durch eine ANOVA mit Messwiederholung zeigt ähnliche Ergebnisse (allerdings ohne

Berücksichtigung der Lehrkraft als zufällige Variable), $F(2,821) = .489$, $p = n. s.$

Der zufällige Faktor „Lehrkraft“ wird dagegen signifikant, $F(22,821) = 2.742$, $p = .038$, $\eta^2 = .003$. Die Effektstärke liegt allerdings im niedrigen Bereich.

Die Wechselwirkung zwischen beiden Faktoren wird hoch signifikant, $F(12,821) = 2.745$, $p = .001$, $\eta^2 = .003$. Hierbei handelt es sich jedoch ebenfalls um einen kleinen Effekt.

5.2.1. Entwicklung des Fachwissens bei leistungsschwachen Schülerinnen und Schülern

Größere Unterschiede können beobachtet werden, wenn die Entwicklung von Schülerinnen und Schülern mit wenig Vorwissen betrachtet wird. Aufgrund der Leistungen im Pretest wurden sämtliche Schülerinnen und Schüler in drei Gruppen eingeteilt (Terzile). Der deskriptive Blick auf die Unterschiede des Lernzuwachses lässt vermuten, dass diese leistungsschwächere Schülergruppe stärker von detaillierten Anleitungen profitiert – der Lernfortschritt ist in Vergleichsgruppe C am geringsten (Abb. 4).

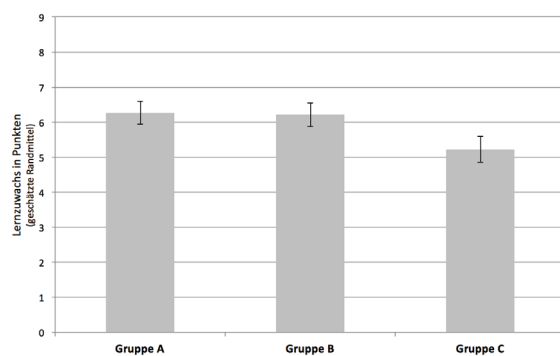


Abb. 4: Schülerinnen und Schüler mit geringem Vorwissen – Lernzuwachs differenziert nach verschiedenen Experimentiersituationen

Allerdings können die Unterschiede zwischen den Vergleichsgruppen lediglich mit der Wechselwirkung zwischen der Experimentiersituation und der Lehrkraft erklärt werden. Weder der feste Faktor ($F(2,299) = .809$, $p = n. s.$) noch der zufällige Faktor allein werden signifikant, $F(22,299) = .962$, $p = n. s.$ Betrachtet man hingegen die Wechselwirkung, kann ein hochsignifikanter Effekt festgestellt werden, $F(12,299) = 3.182$, $p < .001$, $\eta^2 = .011$. Erneut weist die Berechnung der Effektstärke auf einen kleinen Effekt hin.

5.3. Aktuelles Interesse an Physik

Zur Erhebung des aktuellen Interesses an Physik wurden die Schülerinnen und Schüler gebeten, ihr Interesse in Bezug auf die vorangegangenen Unterrichtsstunden auf einer Skala von eins (niedrig) bis fünf (hoch) zu bewerten. Daher wurde diese Skala nur während der Posttests zum Einsatz gebracht.

Der deskriptive Blick auf die Daten (Abb. 5) legt nahe, dass beide Methoden des Schülerexperiments größeres Interesse an Physik hervorrufen als der

Unterricht mit Demonstrationsexperimenten (Gruppe A). Es sei aber gleichwohl darauf hingewiesen, dass auch die „Demo“-Gruppe im Mittel ihr Interesse an Physik knapp oberhalb des Wertes drei bewertet hat. Damit liegt diese Schülergruppe ebenfalls, wenn auch nicht so deutlich, im „positiven“ Bereich der Interessenseinschätzung.

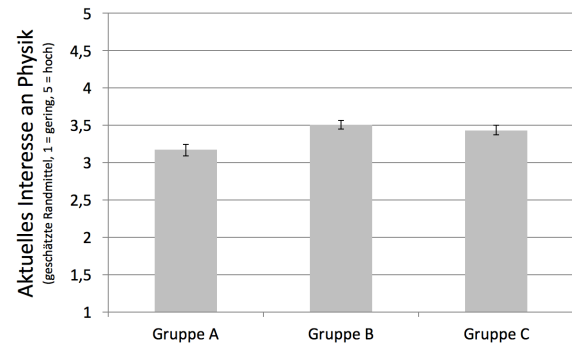


Abb. 5: Das aktuelle Interesse an Physik der Schülerinnen und Schüler, erhoben direkt im Anschluss an die Intervention. Differenziert nach den verschiedenen Experimentiersituationen.

Die Varianzanalyse mittels ANOVA liefert folgende Ergebnisse. Der feste Faktor „Experimentiersituation“ wird nicht signifikant, $F(2,411) = .764$, $p = n. s.$ Auch der zufällige Faktor „Lehrkraft“ wird nicht signifikant, $F(13,411) = .554$, $p = n. s.$ Die Wechselwirkung zwischen beiden Faktoren wird dagegen hoch signifikant, $F(3,411) = 4.476$, $p = .004$, $\eta^2 = .01$, mit einem kleinen Effekt.

Die Unterschiede im Interesse an Physik zwischen den drei Gruppen lassen sich weder allein auf einen Effekt des festen Faktors „Treatment“, noch allein auf einen Effekt des zufälligen Faktors „Lehrkraft“ zurückführen. Die Wechselwirkung zwischen beiden Faktoren wird dagegen höchst signifikant. Die Effektstärke des Wechselwirkungsfaktors liegt allerdings mit $\eta^2 = .01$ im niedrigen Bereich.

6. Diskussion und Einschränkungen der Studie

Mit der vorgestellten Untersuchung konnte nachgewiesen werden, dass der vorstrukturierte und von den teilnehmenden Lehrkräften durchgeführte Physikunterricht äußerst lernwirksam war. Während der Intervention lernten die Schülerinnen und Schüler signifikant hinzu (verbunden mit einer großen Effektstärke).

Die Hypothese, dass Schülerexperimente sowohl für die Entwicklung des Fachwissens als auch für das aktuelle Interesse an Physik eher geeignet sind, kann allerdings nicht bestätigt werden. Man könnte vermuten, dass dieses Nullergebnis dadurch erzeugt wurde, dass das Design der Studie kein ausreichend scharfes Messen zuließ. Der gefundene Effekt ist tatsächlich gering. Allerdings konnte in der vorliegenden Studie die Nullhypothese nicht nur nicht abgelehnt, sondern darüber hinaus bestätigt werden. Mit Bezug auf die dargestellten Daten konnten wir

einen Betafehler von .022 nachweisen. Um die Nullhypothese als bestätigt annehmen zu dürfen, bedarf es einer ausreichenden Teststärke. Diese ist mit $1 - \beta = 0.97$ gegeben.

Die Bewertung dieses Ergebnisses muss im Kontext weiterer Unterrichtsziele gesehen werden. Wenn sich aus Sicht der Lehrkraft hierbei der Einsatz von Schülerexperimenten anbietet, so ergibt sich offenbar für das Erreichen der fachlichen Ziele hieraus kein Nachteil: Trotz der höheren Komplexität von Schülerexperimenten schneiden Schülerinnen und Schüler dieser Experimentiersituation nicht signifikant schlechter ab als solche, die Demonstrationsexperimente beobachtet haben.

Darüber hinaus weist die Auswertung der erhobenen Daten auf die Wechselwirkung zwischen den verglichenen Experimentiersituationen und den unterrichtenden Lehrkräften hin. Mit aller gebotenen Vorsicht bei der Interpretation dieser Wechselwirkung (nur 15 der insgesamt 23 Lehrkräfte unterrichteten Parallelklassen nach zwei verschiedenen Experimentiersituationen), bedeutet die signifikante Wechselwirkung mit einem kleinen Effekt, dass es einen Einfluss hat, welche Experimentiersituation bei welcher Lehrkraft erlebt wurde.

Ähnliche Ergebnisse finden auch Schwichow, Zimmerman, Croker & Härtig (2016) in ihrer Analyse des allgemeinen Lernfortschritts durch den Einsatz von Experimenten im Physikunterricht. Sie schlussfolgern, dass nicht das Experimentieren allein ausschlaggebend für den Lernerfolg ist, sondern ein besonderes Augenmerk auf die Aufgabenstellung gelegt werden sollte. In einer Folgestudie untersuchen wir dagegen, ob sich die gefundene Wechselwirkung zwischen der unterrichtenden Lehrkraft und der jeweiligen Experimentiersituation durch die Überzeugungen der Lehrkräfte erklären lassen. Hierfür betrachten wir die Überzeugungen von Lehrkräften bezüglich der Physik als Unterrichtsfach sowie als Wissenschaft.

Die Ergebnisse der vorgestellten Studie sprechen dafür, dass eine bestimmte Art des Experimentierens im Physikunterricht nicht einer anderen vorzuziehen ist und sind unseres Erachtens somit ein Beleg für den Nutzen einer Vielfalt von Unterrichtssituationen.

7. Literatur

- Anders, C., Berg, R., Christina, V., Bergendahl, B., Bruno, K., & Lundberg, S. (2003). Benefiting from an open-ended experiment? A comparison of attitudes to, and outcomes of, an expository versus open-inquiry version of the same experiment. *International Journal of Science Education*, 25(3), 351-372.
- Atkinson, E. P. (1981). Influence of practical work on test performance. *Research in Science Education*, 11(1) 87-93.
- Behrendt, H. (1991). *Physikalische Schulversuche: Didaktische Theorie, methodische Praxis und die Einstellung von Schülern zur Auswahl der Versuchsgeräte*. Kiel, Germany: IPN.
- Bohl, T. (2001). Wie verbreitet sind offene Unterrichtsmethoden? *Pädagogische Rundschau*, 55, 271-287.
- Börlin, J. (2012). Das Experiment als Lerngelegenheit: Vom interkulturellen Vergleich des Physikunterrichts zu Merkmalen seiner Qualität. In H. Niedderer, H. Fischler & E. Sumfleth (Hrsg.). *Studien zum Physik- und Chemielernen* (Vol. 132). Berlin, Germany: Logos.
- Bortz, J., & Döring, N. (2006). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler*. Heidelberg, Germany: Springer.
- Etkina, E., Van Heuvelen, A., Brookes, D. T., & Mills, D. (2002). Role of Experiments in Physics Instruction – A Process Approach. *The Physics Teacher*, 40, 351-355.
- Field, A. (2011). *Discovering statistics using spss: And sex and drugs and rock 'n' roll*. London, Great Britain: Sage.
- Garrett, M. S., & Roberts, I. F. (1982). Demonstration versus small group practical work in science education: a critical review of studies since 1900. *Studies in Science Education*, 9, 109-146.
- Hacking, I. (1983). *Representing and intervening: Introductory topics in the philosophy of natural science*. Cambridge, MA: Harvard University Press.
- Haladyna, T. M. (1994). *Developing and validating multiple-choice test items*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Hattie, J. (2013). *Lernen sichtbar machen*. Schneider Verlag Hohengehren.
- Heller, K. A., & Perleth, C. (2000). *Kognitiver Fähigkeitstest für 4. bis 12. Klassen, Revision (KFT 4-12+R)*. Göttingen, Germany: Beltz.
- Hodson, D. (1993). Re-thinking old ways: Towards a more critical approach to practical work in school science. *Studies in Science Education*, 22, 85-142.
- Hoffmann, L., Häußler, P., & Lehrke, M. (1998). *Die IPN-Interessenstudie Physik*. Kiel, Germany: IPN.
- Hofstein, A., & Lunetta, V. N. (1982). The role of the laboratory in science teaching: Neglected aspects of research. *Review of Educational Research*, 52, 201-217.
- Hofstein, A., & Lunetta, V. N. (2004). The Laboratory in Science Education: Foundations for the Twenty-First Century. *Science Education*, 88, 28-54.
- Hofstein, A., & Mamlok-Naaman, R. (2007). The laboratory in science education: The state of the art. *Chemistry Education Research and Practice*, 8(2), 105-107.
- Hopf, M. (2007). Problemorientierte Schülerexperimente. In H. Niedderer, H. Fischler, E. Sum-

- fleth (Hrsg.). *Studien zum Physik- und Chemielernen*. Band 68. Berlin: Logos.
- Höttecke, D., & Rieß, F. (2015). Naturwissenschaftliches Experimentieren im Lichte der jüngeren Wissenschaftsforschung – Auf der Suche nach einem authentischen Experimentbegriff der Fachdidaktik. *ZfDN* 21, 127-139.
- Karaböcek, F. & Erb, R. (2014). Typische Experimente im Physikunterricht. In S. Bernholt (Hrsg.), *Naturwissenschaftliche Bildung zwischen Science- und Fachunterricht. Gesellschaft für Didaktik der Chemie und Physik, Jahrestagung in München 2013*, S. 450 - 452. Kiel: IPN.
- Karaböcek, F. & Erb, R. (2015). Survey Experimente – Der Einsatz von Experimenten im Physikunterricht. In: S. Bernholt (Hrsg.), *Heterogenität und Diversität - Vielfalt der Voraussetzungen im naturwissenschaftlichen Unterricht. Gesellschaft für Didaktik der Chemie und Physik, Jahrestagung in Bremen 2014*, S. 399-401. Kiel: IPN.
- Klahr, D., Zimmerman, C., & Jirout, J. (2011). Educational interventions to advance children's scientific thinking. *Science*, 333, 971-975.
- Klieme, E., Rakoczy, K. (2008). Empirische Unterrichtsforschung und Fachdidaktik. Outcomeorientierte Messung und Prozessqualität des Unterrichts. *Zeitschrift für Pädagogik* 54 2, 222-237
- Lienert G., & Raatz, U (1998). *Testaufbau und Testanalyse*. Weinheim, Germany: Psychologie Verlag.
- Lunetta, V. N., Hofstein, A., & Clough, M. P. (2007). Teaching and learning in the school science laboratory: An analysis of research, theory, and practice. In S. K. Abell & N. G. Lederman (Eds.), *Handbook of research on science education* (pp. 393-431). Mahwah NJ: Lawrence Erlbaum.
- Lynch, P. P., & Ndyetabura, V. L. (1983). Practical work in schools: An examination of teachers' stated aims and the influence of practical work according to students. *Journal of Research in Science Teaching*, 20, 623-632.
- Millar, R. (2010). Practical work. In J. Osborne & J. Dillon (Eds.). *Good practice in science teaching: What research has to say* (pp. 108-134). Maidenhead, Great Britain: Open University Press.
- Muth, L. & Erb, R. (2017). Einfluss der Auswertephase von Experimenten im Physikunterricht. In C. Maurer (Hrsg.), *Implementation fachdidaktischer Innovation im Spiegel von Forschung und Praxis. Gesellschaft für Didaktik der Chemie und Physik, Jahrestagung in Zürich 2016*, S. 296. Universität Regensburg
- Muth, L. (2018). Einfluss der Auswertephase von Experimenten im Physikunterricht. Ergebnisse einer Interventionsstudie zum Zuwachs von Fachwissen und experimenteller Kompetenz von Schülerinnen und Schülern. In H. Niedderer, H. Fischler, E. Sumfleth (Hrsg.), *Studien zum Physik- und Chemielernen, Band 252*, Berlin: Logos Verlag.
- Müller, K., Prenzel, M., Seidel, T., Schiepe-Tiska, A., & Kjærnsli, M. (2016). Science Teaching and Learning in Schools: Theoretical and Empirical Foundations for Investigating Classroom-Level Processes. In S. Kuger, E. Klieme, N. Jude, & D. Kaplan (Hrsg.), *Assessing context of learning world-wide – Extended context assessment framework and documentation of questionnaire material* (S. 423-446). New York: Springer.
- National Endowment for Science, Technology and the Arts (NESTA) (2008). *Science Teachers Survey* [.pdf]. London, Great Britain: NESTA. Retrieved from: <https://www.nfer.ac.uk/what-we-offer/teacher-voice/PDFs/science-enquiry.pdf>
- Oplt, J. (2013). *Validierung eines Testinstruments in Optik*. (Examensarbeit). Institut für Didaktik der Physik, Goethe-Universität, Frankfurt.
- Osborne, J., Simon, S., & Collins, S. (2003). Attitudes towards science: A review of the literature and its implications. *International Journal of Science Education*, 25, 1049-1079.
- Potvin, P., & Hasni, A. (2014). Interest, motivation and attitude towards science and technology at K-12 levels: a systematic review of 12 years of educational research. *Studies in Science Education*, 50(1), 85-129.
- Potvin, P., Hasni, A., & Sy, O. (2017). Using inquiry-based interventions to improve secondary students' interest in science and technology. *European Journal of Science and Mathematics Education* 5(3), 262-270.
- Rönnebeck, S., Bernholt, S., & Ropohl, M. (2016). Searching for a common ground: A literature review of empirical research on scientific inquiry activities. *Studies in Science Education*, 52(2), 161-197.
- Rumann, S. (2005). Kooperatives Experimentieren im Chemieunterricht: Entwicklung und Evaluation einer Interventionsstudie zur Säure-Base-Thematik. In H. Niedderer, H. Fischler & E. Sumfleth (Hrsg.). *Studien zum Physik- und Chemielernen* (Vol. 45). Berlin, Germany: Logos.
- Schecker, H., Neumann, K., Theyßen, H., Eickhorst, B., & Dickmann, M. (2016). Stufen experimenteller Kompetenz. *Zeitschrift für Didaktik der Naturwissenschaften*, 22(1), 197-213.
- Schreiber, N., Theyßen, H., & Schecker, H. (2009). Experimentelle Kompetenz messen?! *Physik und Didaktik in Schule und Hochschule*, 3(8), 92-101.
- Singer, S. R., Hilton, M. L., & Schweingruber, H. A. (2006). *America's Lab Report. Investigations in*

- High School Science*. Washington, DC: The National Academies.
- Schulz, A. (2011). Experimentierspezifische Qualitätsmerkmale im Chemieunterricht: Eine Videostudie. In H. Niedderer, H. Fischler & E. Sumfleth (Hrsg.), *Studien zum Physik- und Chemielernen* (Vol. 113). Berlin, Germany: Logos.
- Schwichow, M., Zimmerman, C., Croker, S., & Härtig, H. (2016). What students learn from hands-on activities. *Journal of Research in Science Teaching*, 1-23.
- Shulman, L. S., & Tamir, P. (1973). Research on teaching in the natural sciences. In R. M. W. Travers (Ed.), *Second handbook of research on teaching* (pp. 1098-1148). Chicago, IL: Rand-McNally.
- Tabachnick, B. G., & Fidell, L. S. (2007). *Experimental designs using SPSS*. Belmont, CA: Thomson Higher Education.
- Tesch, M. (2005). Das Experiment im Physikunterricht: Didaktische Konzepte und Ergebnisse einer Videostudie [Practical work in school Physics: Didactically concepts and results of a video study]. In H. Niedderer, H. Fischler & E. Sumfleth (Hrsg.), *Studien zum Physik- und Chemielernen* (Vol. 42). Berlin, Germany: Logos.
- Theyßen, H., Schecker, H., Neumann, K., Eickhorst, B., & Dickmann, M. (2016). Messung experimenteller Kompetenz – ein computergestützter Experimentierertest. *PhyDid A – Physik und Didaktik in Schule und Hochschule*, 1(15), 26-48.
- Thijs, G. D., & Bosch, G. M. (1995). Cognitive effects of science experiments focusing on students' preconceptions of force: A comparison of demonstration and small-group practicals. *International Journal of Science Education*, 17(3), 311-323.
- Thompson, J. & Soyibo, K. (2002). Effects of lecture, teacher demonstration, discussion and practical work on 10th graders' attitudes to chemistry and understanding of electrolysis. *Research in science and technological education*, 20 (1), 25-35.
- Watson, J. R., Prieto, T., & Dillon, J. (1995). The effect of practical work on students' understanding of combustion. *Journal of Research in Science Teaching*, 32(5), 487-502.
- Welzel, M., Haller, K., Bandiera, M., Hammelev, D., Koumaras, P., Niedderer, H., ... von Aufschneider, S. (1998). Ziele, die Lehrende mit dem Experimentieren in der naturwissenschaftlichen Ausbildung verbinden: Ergebnisse einer europäischen Umfrage. *Zeitschrift für Didaktik der Naturwissenschaften*, 4(1), 29-44.
- Winkelmann, J. & Erb, R. (2011). Fachwissenszuwachs durch Schüler- und Lehrerexperimente im gymnasialen Physikunterricht. In H. Grötzebach & V. Nordmeier (Hrsg.), *PhyDid B – Didaktik der Physik: Beiträge zur DPG-Frühjahrstagung des Fachverbands Didaktik der Physik in Münster 2011*, Beitrag 15.02.
- Winkelmann, J. & Erb, R. (2014). Lernzuwachs durch Schüler- und Demonstrationsexperimente: Experimentiervorschläge zur geometrischen Optik. *Mathematisch- Naturwissenschaftlicher Unterricht*, 67 (7), 394-401.
- Winkelmann, J. (2015). Auswirkungen auf den Fachwissenszuwachs und auf affektive Schülermerkmale durch Schüler- und Demonstrationsexperimente im Physikunterricht. In H. Niedderer, H. Fischler, E. Sumfleth (Hrsg.), *Studien zum Physik- und Chemielernen. Band 179*. Berlin: Logos Verlag.
- White, R. T. (1996). The link between the laboratory and learning. *International Journal of Science Education*, 18(7), 761-774.