

Jeremias Weber  
 S. Franziska C. Wenzel  
 Jan Winkelmann  
 Mark Ullrich  
 Roger Erb  
 Holger Horz

Goethe-Universität Frankfurt

## **Veränderung von Fachwissen in verschiedenen Experimentiersituationen**

### **Motivation**

Experimente, insbesondere Schülerexperimente machen einen großen Anteil von naturwissenschaftlichem Unterricht aus (Duit & Wodzinski, 2010), wobei hauptsächlich sogenannte „Kochbuch“-Experimente („cook-book“ list of tasks“, Hofstein & Lunetta, 2004, S. 47) Anwendung finden. Verschiedene Autoren fordern allerdings offenere Lernsituationen, die zwar mehr kognitive Anstrengungen der Lernenden erfordern, sich allerdings lernförderlich auswirken können und sehr geeignet für die Arbeit in heterogenen Lerngruppen scheinen (Bunterm et al., 2014; Koksal & Berberoglu, 2014; Hofstein & Lunetta, 2004; Wodzinski et al., 2007). Unklar bleibt jedoch noch (Hofstein & Lunetta 2004, Lazonder & Harmsen 2016), ob eine bestimmte Art oder ein spezifischer Umfang der Anleitung einen klaren Vorteil mit sich bringt. Dazu passend zeigen Ergebnisse einer Studie von Winkelmann (2015) zur Wirkung von Experimentiersituationen auf das Fachwissen von Schülerinnen und Schülern keinen signifikanten Unterschied zwischen Demonstrations- oder Schülerexperimenten (mit offenen oder angeleiteten Aufgabenstellungen). Allerdings konnte eine signifikante, kleine Wechselwirkung zwischen der durchführenden Lehrkraft und der Experimentiersituation detektiert werden.

### **Studie zur Kompetenzmessung und Kompetenzförderung in leistungsheterogenen Lerngruppen im experimentierbasierten Physikunterricht (KoPhy-Studie)**

#### *Forschungsfragen und Studiendesign*

Folgende drei Forschungsfragen wurden formuliert:

- 1.1 Wie wirken sich die bei Winkelmann (2015) unterschiedenen Experimentiersituationen im Physikunterricht auf die Entwicklung in den Bereichen „Fachwissen“ und „Erkenntnisgewinnung“ und auf das aktuelle Interesse der Schülerinnen und Schüler an Physik aus?
- 1.2 Welche Unterschiede zeigen sich in heterogenen Leistungsgruppen aufgrund der unterschiedlichen Experimentiersituationen im Physikunterricht in Bezug auf die Entwicklung in den Bereichen „Fachwissen“ und „Erkenntnisgewinnung“?
2. Welche Auswirkungen hat die Interaktion von Lehrercharakteristika und Experimentiersituation auf die Kompetenzentwicklung von Schülerinnen und Schülern im Fach Physik?

Diese Forschungsfragen sollen in unserer, vom BMBF geförderten, Interventionsstudie beantwortet werden. Dafür werden die bei Winkelmann (2015) vorgeschlagenen Unterrichtsreihen mit Fokus auf die verschiedenen Experimentiersituationen verwendet (vgl. Weber et al., 2016). In der Studie werden die Teilnehmenden (Schülerinnen und Schüler sowie Lehrkräfte) zu drei verschiedenen Testzeitpunkten befragt: Ein Pretest direkt vor und ein Posttest direkt nach der Unterrichtsreihe sowie ein Follow-Up-Tests 4 bis 12 Wochen später.

Für die Follow-Up-Erhebung wurde ein Planned-Missing-Design (Little & Rhemtulla, 2013) genutzt. Dabei nehmen die Teilnehmenden immer nur an einem der drei Follow-Up-

Messzeitpunkte (4, 8, oder 12 Wochen später) teil. Die angepeilte Stichprobengröße lässt dieses Vorgehen bei hoher zu erwartender Teststärke zu.

#### *Messinstrumente*

Im Pretest werden kognitive Leistungsfähigkeit (KFT-V3, -N2, Heller & Perleth, 2000) und personenbezogene Daten der Schülerinnen und Schüler erhoben. Die Lehrkräfte werden gleichzeitig zu ihren Überzeugungen zum Physikunterricht und zur Physik als Wissenschaft befragt (Lamprecht, 2011). Im Post-Test wird das aktuelle Interesse der Schülerinnen und Schüler an Physik mit einer an den Physikunterricht adaptierten Skala von Schulz (2011) erhoben. Der Stand in den Bereichen Fachwissen und Erkenntnisgewinnung wird zu allen Testzeitpunkten gemessen. Um Veränderung im Bereich Erkenntnisgewinnung zu messen, werden aus dem Test zur prozessbezogenen naturwissenschaftlichen Grundbildung von Glug (2009) Itemsets verwendet. Diese wurden bereits unter Nutzung von Modellen der Item-Response-Theorie (IRT, van der Linden & Hambleton, 2013) evaluiert. Für den Bereich Fachwissen wurde von uns ein ebenfalls IRT-basierter Test entwickelt.

#### *Fachwissenstest: Konzeption und Implementierung*

Zunächst wurden die von Winkelmann (2015) eingesetzten Items IRT-skaliert und hinsichtlich zentraler psychometrischer Aspekte analysiert. Aus den 54 Items von Winkelmann konnten 32 Items selektiert werden. Ein Teil der eliminierten Items wurde überarbeitet und zusätzlich 30 neue Items in einer Expertengruppe entwickelt. Zusammen mit den bereits vorhandenen Items wurden diese anschließend im Zuge einer Pilotstudie kalibriert. An der Pilotstudie nahmen 301 Personen teil, sowohl Schülerinnen und Schüler als auch Lehramtsstudierende der Universitäten Köln und Frankfurt. Wiederum wurden die Items anhand der Pilotierungsdaten IRT-skaliert, auf psychometrische Güte geprüft und gegebenenfalls für den Fachwissenstest selektiert. Aus dem so gewonnenen Itempool zur Messung des Fachwissens wurden Testhefte für alle drei Messzeitpunkte zusammengestellt. Jedes messzeitpunktspezifische Testheft bestand aus 16 Items, mit einer Zahl von 4 bis 8 Items die sich mit Testheften anderer Messzeitpunkte überlappen. Auf diese Weise ist eine Verknüpfung der Resultate trotz unterschiedlicher dargebotener Items über die Messzeitpunkte möglich. Um die zu erwartende Personenfähigkeit am jeweiligen Messzeitpunkt möglichst genau zu erfassen wurden im Pretest eher einfache Items genutzt, im Posttest dafür schwerere. Zudem kam ein balanciertes unvollständiges Testheftdesign (Osterlind & Everson, 2009) zum Einsatz, um Reihenfolgeeffekte zu vermeiden.

#### **Erste Erkenntnisse der Hauptstudie**

Nach dem ersten Datencleaning können zum jetzigen Zeitpunkt die Angaben von 44 Lehrkräften und 1094 Schülerinnen und Schülern in die nachfolgenden Analysen einbezogen werden. Die durchschnittliche Klassengröße betrug 25 Schülerinnen und Schüler. Hauptsächlich waren Klassen der 7. Jahrgangsstufe (73%) vertreten, es gab aber auch Klassen der 8. Jahrgangsstufe (27%). Schüler machten 44% der Stichprobe aus, 46% waren Schülerinnen. Die übrigen Teilnehmer gaben kein Geschlecht an.

#### *Veränderung des Fachwissens*

Der Zuwachs des, durch den zuvor entwickelten und vorgestellten Test ermittelten Fachwissens von Pretest zu Posttest ist unabhängig von der Experimentiersituation signifikant und sehr groß (partiell  $\eta^2 = 0,365$ ). Es lässt sich kein Unterschied zwischen den verschiedenen Experimentiersituationen finden. Nach Ende der Unterrichtsreihe sinkt das Fachwissen der Schülerinnen und Schüler erwartungsgemäß wieder ab, allerdings ist dies in der Kochbuch-Experimentiersituation signifikant weniger ausgeprägt, wobei der detektierte Effekt sehr klein ausfällt (partiell  $\eta^2 < 0,01$ ).

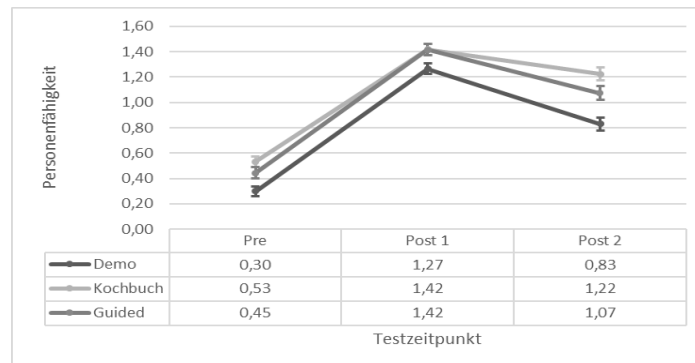


Abb. 1: Personenfähigkeit im Bereich Fachwissen

Nach Eliminierung nicht vollständiger Datensätze betrug die Stichprobengröße hier 867 Schülerinnen und Schüler (Demo: 302, Kochbuch: 291, Guided 274).

#### Aktuelles Interesse, abhängig von Geschlecht und Experimentiersituation

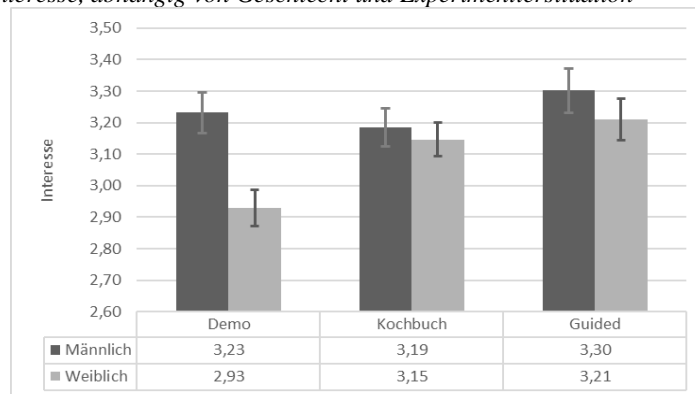


Abb. 2: Aktuelles Interesse an der Unterrichtsreihe

Das aktuelle Interesse wurde auf einer Likert-Skala angegeben (1: nicht interessiert, 5: interessiert). Während sich bei den Schülern keine signifikante Wechselwirkung mit der Experimentiersituation ergibt und Schüler und Schülerinnen keine signifikanten Unterschiede bei den Schülerexperimenten zeigen, haben Schülerinnen ein deutlich geringeres aktuelles Interesse an der Demo-Experimentiersituation ( $d_{\text{Cohen}} = 0,38$ ;  $p < 0,5$ ).

#### Fazit

Der neu entwickelte Fachwissenstest hat sich als gut verwendbar erwiesen. Während die Daten, insbesondere im Hinblick auf Heterogenität, Kompetenzbereich Erkenntnisgewinnung und Einfluss der Überzeugungen der Lehrkräfte noch anhand weiterer Analysen und unter Berücksichtigung der Mehrebenenstruktur untersucht werden müssen, zeigt sich bereits, dass Befunde vorheriger Studien bestätigt werden können: Die verschiedenen Experimentiersituationen bieten jeweils keine deutlichen Vor- noch Nachteile bei der Erlangung von Fachwissen. Dagegen zeigt sich im Bereich des aktuellen Interesses ein Hinweis darauf, dass insbesondere Schülerinnen wohl weniger stark durch Demonstrationsexperimente angesprochen werden als durch Schülerexperimente. Zieht man den starken Zusammenhang von Interesse, Motivation und Selbstwirksamkeitserwartung in die Betrachtung mit ein und bedenkt zudem, dass im deutschen Schulsystem viele Schülerinnen und Schüler im Laufe ihrer Schulzeit ein ungünstiges Einstellungsmuster bzw. Selbstkonzept gegenüber naturwissenschaftlichen Fächern und Berufen entwickeln, ist dieses Ergebnis umso beachtenswerter.

**Literatur**

- Bunterm, T., Lee, K., Ng Lan Kong, J., Srikoon S., Vangpoomyai, P., Rattanavongsa, J., & Rachahoon, G. (2014). Do Different Levels of Inquiry Lead to Different Learning Outcomes? A comparison between guided and structured inquiry. *International Journal of Science Education*, 36(12), 1937-1959, DOI:10.1080/09500693.2014.886347.
- Duit, R., & Wodzinski, C. T. (2010). Merkmale guten Physikunterrichts. In: R. Duit (Hrsg.). Piko-Briefe. Der fachdidaktische Forschungsstand kurzgefasst. IPN Kiel. Abgerufen von: <http://www.ipn.uni-kiel.de/de/das-ipn/abteilungen/didaktik-der-physik/piko/pikobriefe032010.pdf>
- Glug, I. (2009). Entwicklung und Validierung eines Multiple-Choice-Tests zur Erfassung prozessbezogener naturwissenschaftlicher Grundbildung. Kiel: IPN.
- Hofstein, A., & Lunetta, V. N. (2004). The Laboratory in Science Education: Foundations for the Twenty-First Century. *Science Education*, 88, 28-54.
- Heller, K. A., & Perleth, C. (2000). Kognitiver Fähigkeitstest für 4. bis 12. Klassen, Revision (KFT 4-12+R). Göttingen: Beltz Test GmbH.
- Koksal, E. A., & Berberoglu, G. (2014). The Effect of Guided-Inquiry Instruction on 6th Grade Turkish Students' Achievement, Science Process Skills, and Attitudes Toward Science. *International Journal of Science Education*, 36(1), 66-78, DOI: 10.1080/09500693.2012.721942.
- Lamprecht, J. (2011). Ausbildungswege und Komponenten professioneller Handlungskompetenz. Vergleich von Quereinsteigern mit Lehramtsabsolventen für Gymnasien im Fach Physik. In H. Niedderer, H. Fischler & E. Sumfleth (Hrsg.). *Studien zum Physik- und Chemielernen*. Band 125. Berlin: Logos Verlag.
- Lazonder, A. W., & Harmsen, R. (2016). Meta-analysis of inquiry-based learning: Effects of guidance. *Review of Educational Research*, 86, 681–718. <https://doi.org/10.3102/0034654315627366>.
- Little, T. D., & Rhemtulla, M. (2013). Planned Missing Data Designs for Developmental Researchers. *Child Development Perspectives*, 7, 199–204. doi:10.1111/cdep.12043.
- Osterlind, S. J., & Everson, H. T. (2009). *Differential item functioning* (Vol. 161). Sage Publications.
- Schulz, A. (2011). Experimentierspezifische Qualitätsmerkmale im Chemieunterricht: Eine Videostudie. In: H. Niedderer, H. Fischler, & E. Sumfleth (Hrsg.). *Studien zum Physik und Chemielernen*. Band 113. Berlin: Logos Verlag.
- van der Linden, W. J., & Hambleton, R. K. (Hrsg.). (2013). *Handbook of modern item response theory*. Springer Science & Business Media.
- 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.
- Weber, J., Winkelmann, J., Erb, R., Wenzel, S. F. C., Ullrich, M., & Horz, H. (2016). Entwicklung von Messinstrumenten zum Kompetenzzuwachs anhand von Modellen der IRT. In: *PhyDid B – Beiträge zur DPG-Frühjahrstagung des Fachverbandes Didaktik der Physik in Hannover*.
- Wodzinski, R., Wodzinski, C.T., & Hepp, R. (Hrsg.) (2007). Themenheft "Differenzierung im Physikunterricht". *Unterricht Physik* Nr. 99/100. Seelze: Friedrich Verlag.