

Albert Teichrew<sup>1</sup>  
Roger Erb<sup>1</sup>

<sup>1</sup>Goethe-Universität Frankfurt am Main

## **Online-Förderung physikalischer Konzepte mit Modellen und Experimenten**

In diesem Beitrag werden Ergebnisse einer Begleitstudie zu einem Online-Kurs für Schüler\*innen zur Optik vorgestellt und diskutiert. In dem Gesamtprojekt wird untersucht, unter welchen Bedingungen die Arbeit mit dynamischen Modellen und hypothesenüberprüfenden Experimenten einen Einfluss auf Schüler\*innenvorstellungen hat.

### **Hintergrund**

Das Experimentieren ist eine beliebte Unterrichtsmethode (Tesch & Duit, 2004). Allerdings zeigen Unterrichtsbeobachtungen, dass sich Lernende dabei wenig Gedanken über den physikalischen Hintergrund ihrer experimentellen Tätigkeit machen (Abrahams & Millar, 2008). Vor allem Lehrkräfte verbinden mit dem Experimentieren unter anderem auch den Ausbau von Fachwissen (Karaböcek & Erb, 2014). Didaktiker\*innen schätzen Fachwissen hingegen als Voraussetzung für ein sinnstiftendes Experimentieren ein (Härtig, Neumann, & Erb, 2017). Als möglichen Lösungsansatz arbeiten wir deshalb mit einer Unterrichtsstruktur, die Modelle physikalischer Phänomene für die Suche im Hypothesenraum (Scientific Discovery as Dual Search-Modell nach Klahr & Dunbar, 1988) und hypothesenüberprüfende Experimente in einem gemeinsamen Erkenntnisprozess verbindet (Teichrew & Erb, 2018).

Als Werkzeug der Modellierung verwenden wir die Dynamische Geometrie-Software (DGS) *GeoGebra* (geogebra.org). Die Besonderheit einer DGS ist, dass sich Veränderungen an einem Objekt auch auf alle verknüpften Objekte auswirken. Dadurch wird ein virtuelles Experimentieren mit dynamischen Modellen möglich (Erb, 2017), um Vorhersagen für das reale Experiment zu gewinnen (Teichrew & Erb, 2019b). Wendet man den Modellbegriff nach Mahr (2008) auf dynamische Modelle an, handelt es sich um geometrische Visualisierungen und interaktive Simulationen (Modell als Objekt), die beim Aufbau mentaler Modelle bezüglich physikalischer Phänomene helfen (Modell von etwas) und im Experiment überprüft werden können (Modell für etwas).

Effekte beim Lernen mit (GeoGebra-)Modellen sind dann zu erwarten, wenn die wahrgenommene Information im Arbeitsgedächtnis verarbeitet und mit dem Vorwissen verknüpft wird (Cognitive Theory of Multimedia Learning nach Mayer, 2005). Die erfolgreiche Verarbeitung von Lernmedien kann allerdings durch kognitive Belastung erschwert werden (Cognitive Load Theory nach Sweller, 1988). Es wird unterschieden zwischen der Belastung, die sich auf den zu behandelnden Inhalt selbst (intrinsic cognitive load, ICL) oder auf das Instruktionsdesign zurückführen lässt. Letztere lässt sich unterteilen in für das Lernen unwirksame (extraneous cognitive load, ECL) und wirksame kognitive Belastung (germane cognitive load, GCL).

### **Online-Kurs**

Aufgrund der COVID-19-Pandemie konnten die für ein Schüler\*innenlabor vorbereiteten Lernaktivitäten zum Thema Optik nicht wie geplant stattfinden. Stattdessen wurden die Lernmaterialien und Testinstrumente interessierten Lehrkräften in Form eines Online-Kurses angeboten, den sie daraufhin ihren Schüler\*innen als digitale Lerngelegenheit für Zuhause empfohlen haben. Es wurden insgesamt vier Module mit den Themenbereichen Licht,

Schatten, Spiegel und Farben zusammengestellt, wobei in jedem Modul mindestens drei Themen behandelt werden. Jedes Thema enthielt eine Fragestellung, die mithilfe eines Modells und Experiments beantwortet werden musste. Die Lernenden wurden zunächst über das jeweilige Thema informiert und konnten danach Untersuchungen an einem dynamischen Modell zum Phänomen durchführen. Um neben dem virtuellen Experimentieren mit GeoGebra-Modellen auch den Vergleich mit realen Beobachtungen zu ermöglichen, wurden einzelne Experimentierschritte abfotografiert und zu interaktiven Bildschirmexperimenten zusammengestellt. Beide Lernmedien enthielten Beschreibungen und Anleitungen, um eine eigenständige Verarbeitung der Darstellungen zu ermöglichen.

Es wurde ein zweistufiger Schüler\*innenvorstellungstest im Pre-Post-Design eingesetzt (Teichrow & Erb, 2019a). Neben der Antwort auf eine Frage muss im zweiten Schritt auch eine Erklärung ausgewählt werden (1 Punkt pro Antwort und 2 Punkte pro Erklärung), wobei die Erklärungen auf bekannten Vorstellungen basieren (Haagen-Schützenhöfer & Hopf, 2018). Die Ergebnisse einer Vorabhebung mit diesem Testinstrument in der Zielgruppe wurden im Rahmen einer Didaktischen Rekonstruktion nach Kattmann et al. (1997) genutzt, um die Inhalte für das Schüler\*innenlabor bzw. den Online-Kurs auszuwählen.

Um außerdem den Lernprozess abbilden zu können, wurden die Antworten der Lernenden auf die Fragestellungen zu drei Messzeitpunkten erfasst: Nach der Themenvorstellung, nach dem Modell und nach dem Experiment. Dafür wurden jedes Mal zufällig drei Aussagen zu dem Thema aus einem Itempool mit richtigen und falschen Aussagen ausgewählt, die richtig zugeordnet werden mussten (0 bis 3 richtige Zuordnungen pro Messzeitpunkt in bis zu vier Themen). Dadurch sollten Wiederholungs- und Kombinationseffekte reduziert werden.

Zur Bestimmung der kognitiven Belastung wurden 28 Items verwendet, die auf den Formulierungen bestehender Messinstrumente basieren. Sie beziehen sich auf konkrete Elemente des Online-Kurses (Themen, Fragestellungen, Zuordnungen) und betrachten Modelle und Experimente getrennt voneinander (jeweils Beschreibung, Anleitung, Durchführung). Die Bestimmung der intrinsischen kognitiven Belastung mit einer 5-stufigen Likert-Skala erfolgt über Indikatoren, die die Schwierigkeit des Inhalts betreffen. Bei der extrinsischen kognitiven Belastung werden die Gestaltung des Materials und die damit verbundenen Lernaktivitäten als Indikatoren herangezogen. Bei der Messung der lernwirksamen kognitiven Belastung hingegen wird entweder die Anstrengung beim Lernen oder die Selbsteinschätzung, durch die Lernaktivität etwas verstanden zu haben, verwendet (vgl. Klepsch, Schmitz, & Seufert, 2017; Leppink, Paas, Van der Vleuten, Van Gog, & Van Merriënboer, 2013; Thees et al., 2020). In unsere Auswahl an Items wurden beide Varianten einbezogen. Außerdem gehen Schnotz, Fries und Horz (2009) davon aus, dass eine hohe extrinsische Belastung ebenso durch zu einfach erscheinendes Material vorliegen kann, sodass für diese Items 9-stufige Skalen mit gegensätzlichen Endpunkten gewählt wurden (z. B. zu einfach und zu schwer, vgl. NASA-TASK LOAD INDEX nach Hart, 2006).

### **Analysen**

Es liegen 228 volle Datensätzen von Studierenden und Schüler\*innen vor, die die Module Licht (n = 125), Schatten (n = 45), Spiegel (n = 55) und Farben (n = 3) bearbeitet haben. Die analysierte Stichprobe bezüglich der Schüler\*innenvorstellungen setzt sich aus Schüler\*innen der Klassenstufen 6, 7 und 8 zusammen, die hauptsächlich an hessischen Gymnasien unterrichtet werden und das Modul Licht bearbeitet haben (n = 97). In diesem Modul wurden lediglich Testitems eingesetzt, die mit dem Modul Licht gefördert werden

konnten (7 von 18 mit max. 21 Punkten). Das durchschnittliche Ergebnis war vor der Bearbeitung der Lernumgebung ( $M = 8.18$ ,  $SD = 4.46$ ) niedriger als nach der Bearbeitung ( $M = 9.27$ ,  $SD = 4.90$ ). Es gibt einen signifikanten Unterschied zwischen den Testergebnissen vor und nach der Bearbeitung der Lernumgebung ( $t(96) = 3.34$ ,  $p < .002$ ,  $d = .34$ , kleiner Effekt).

Die durchschnittliche Zahl der richtigen Zuordnungen in den vier Themen des Moduls war nach der Themenvorstellung am niedrigsten ( $M = 1.93$ ,  $SD = .42$ ), nahm nach der Arbeit mit dem Modell ( $M = 2.08$ ,  $SD = .47$ ) und dem Experiment jeweils zu ( $M = 2.12$ ,  $SD = .53$ ). Eine ANOVA mit Messwiederholung (Sphärizität angenommen: Mauchly  $W(2) = .960$ ,  $p = .144$ ) zeigt, dass die Werte mit dem Messzeitpunkt zusammenhängen ( $F(2, 192) = 8.826$ ,  $p < .001$ ,  $\eta^2_{\text{part}} = .08$ , mittlerer Effekt). Bonferroni-korrigierte post-hoc Tests zeigen einen signifikanten Unterschied zwischen dem ersten und dem zweiten Messzeitpunkt ( $p = .007$ ) und zwischen dem ersten und dem dritten Messzeitpunkt ( $p = .001$ ).

Die Stichprobe für eine Hauptkomponentenanalyse bezüglich der 28 Items zur kognitiven Belastung durch Modelle und Experimente besteht aus allen Datensätzen zum Modul Licht ( $n = 125$ ; Fälle/Item  $> 4$ ). Es liegt eine ausreichend hohe Korrelation vor ( $KMO = .853$ ; Bartlett-Test: Chi-Quadrat(378) = 2262.87,  $p < .001$ ). Eine Überprüfung des Scree-Plots rechtfertigte die Extraktion von vier Komponenten, die eine Gesamtvarianz von 63 % aufklären. Die Varimax-rotierte Lösung lieferte ein Ergebnis, bei dem die meisten Items nur auf einer der vier Komponenten hohe Ladungen zeigten. Als Ergebnis wurden die Komponenten *wahrgenommener Lerngewinn* ( $n_1 = 8$ ;  $\alpha_C = .93$ ; perceived learning gain, PLG), *lernrelevante Anstrengung* ( $n_1 = 8$ ;  $\alpha_C = .88$ ; GCL), ICL ( $n_1 = 4$ ;  $\alpha_C = .77$ ) und ECL ( $n_1 = 8$ ;  $\alpha_C = .91$ ) extrahiert, und zwar unabhängig vom Lernmedium.

Da die verwendete ECL-Skala zwei gegensätzliche Ausprägungen misst und beide einen mutmaßlich negativen Einfluss auf das Lernen haben, wurden drei Gruppen gebildet, die das Material unterschiedlich eingeschätzt haben ( $M_{\text{ECL}} \leq 4,5$ : *unterfordernd* ( $n = 37$ );  $4,5 < M_{\text{ECL}} < 5,5$ : *optimal* ( $n = 36$ );  $M_{\text{ECL}} \geq 5,5$ : *überfordernd* ( $n = 24$ )). Der durchschnittliche Lernzuwachs ist in der Gruppe *überfordernd* ( $M = .67$ ,  $SD = 2.35$ ) etwas höher als in der Gruppe *unterfordernd* ( $M = .35$ ,  $SD = 3.16$ ) und am größten in der Gruppe *optimal* ( $M = 2.14$ ,  $SD = 3.57$ ). Eine einfaktorielle ANOVA (Varianzhomogenität angenommen: Levene-Test:  $F(2, 94) = 1.424$ ,  $p = .246$ ) zeigt, dass sich der durchschnittliche Lernzuwachs statistisch signifikant für die verschiedenen Ausprägungen unterscheidet ( $F(2, 84) = 3.225$ ,  $p = .044$ ,  $\eta^2 = .06$ , mittlerer Effekt). Der Tukey post-hoc Test zeigte einen signifikanten Unterschied ( $p = .045$ ) zwischen den Gruppen *unterfordernd* und *optimal*.

### Diskussion und Ausblick

In einem ersten Durchgang konnte eine messbare Förderung physikalischer Konzepte mit Modellen und Experimenten in einem Online-Kurs erzielt werden. Eine differenzierte Analyse der Ergebnisse einzelner Items und Zuordnungen in verschiedenen Themen stellt die Basis für eine Überarbeitung der Materialien dar, sodass in Zukunft größere Effekte erzielt werden könnten. Es konnten keine messbaren Unterschiede der kognitiven Belastung durch Modelle oder Bildschirmexperimente ermittelt werden, was eine Messung von ECL, GCL und PLG in rein virtuellen Formaten ohne Unterscheidung des Mediums nahelegt. Ein Vergleich der GeoGebra-Modelle mit realen Experimenten steht allerdings noch aus. Es konnte ein Unterschied in der Operationalisierung zwischen GCL und PLG gezeigt werden. Außerdem scheint sich der Einfluss von ECL auf den messbaren Lerngewinn durch Über- und vor allem Unterforderung zu bestätigen. Multiple lineare Regressionsmodelle könnten den Einfluss einzelner Arten kognitiver Belastung weiter aufklären.

## Literatur

- Abrahams, I., & Millar, R. (2008). Does Practical Work Really Work? A study of the effectiveness of practical work as a teaching and learning method in school science. *International Journal of Science Education*, 30(14), 1945–1969.
- Erb, R. (2017). *Optik mit GeoGebra*. Berlin: De Gruyter.
- Haagen-Schützenhöfer, C., & Hopf, M. (2018). Schülervorstellungen zur geometrischen Optik. In H. Schecker, T. Wilhelm, M. Hopf, & R. Duit (Hrsg.), *Schülervorstellungen und Physikunterricht* (S. 89–114). Berlin, Heidelberg: Springer.
- Hart, S. G. (2006). NASA-Task Load Index (NASA-TLX); 20 Years Later. In *Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting* (S. 904–908). Santa Monica: HFES.
- Härtig, H., Neumann, K., & Erb, R. (2017). Experimentieren als Interaktion von Situation und Person: Ergebnisse einer Expertenbefragung. *ZfDN*, 23, 71–80.
- Karaböcek, F., & Erb, R. (2014). Funktionale Aspekte des Experiments: Die Sicht der Lehrkraft. *PhyDid B - Didaktik der Physik - Beiträge zur DPG-Frühjahrstagung*.
- Kattmann, U., Duit, R., Gropengießer, H., & Komorek, M. (1997). Das Modell der Didaktischen Rekonstruktion—Ein Rahmen für naturwissenschaftsdidaktische Forschung und Entwicklung. *ZfDN*, 3(3), 3–18.
- Klahr, D., & Dunbar, K. (1988). Dual Space Search During Scientific Reasoning. *Cognitive Science*, 12(1), 1–48.
- Klepsch, M., Schmitz, F., & Seufert, T. (2017). Development and Validation of Two Instruments Measuring Intrinsic, Extraneous, and Germane Cognitive Load. *Frontiers in Psychology*, 8.
- Leppink, J., Paas, F., Van der Vleuten, C. P. M., Van Gog, T., & Van Merriënboer, J. J. G. (2013). Development of an instrument for measuring different types of cognitive load. *Behavior Research Methods*, 45(4), 1058–1072.
- Mahr, B. (2008). Ein Modell des Modellseins – Ein Beitrag zur Aufklärung des Modellbegriffs. In U. Dirks & E. Knobloch (Hrsg.), *Modelle* (S. 187–220). Frankfurt am Main: Peter Lang Verlag.
- Mayer, R. (2005). *Cognitive Theory of Multimedia Learning*. In R. Mayer (Hrsg.), *The Cambridge Handbook of Multimedia Learning* (S. 31–48). Cambridge: Cambridge University Press.
- Schnotz, W., Fries, S., & Horz, H. (2009). Motivational aspects of cognitive load theory. In M. Wosnitza, S. A. Karabenick, A. Efklides, & P. Nenniger (Hrsg.), *Contemporary motivation research: From global to local perspectives* (S. 69–96). Ashland, OH, US: Hogrefe & Huber Publishers.
- Sweller, J. (1988). Cognitive Load During Problem Solving: Effects on Learning. *Cognitive Science*, 12(2), 257–285.
- Teichrew, A., & Erb, R. (2018). Implementierung modellbildender Lernangebote in das physikalische Praktikum. *PhyDid B - Didaktik der Physik - Beiträge zur DPG-Frühjahrstagung*.
- Teichrew, A., & Erb, R. (2019a). Entwicklung und Evaluation eines zweistufigen Testinstruments für Schülervorstellungen zur Anfangsoptik. *PhyDid B - Didaktik der Physik - Beiträge zur DPG-Frühjahrstagung*.
- Teichrew, A., & Erb, R. (2019b). Videobasierte Analyse des Lernens mit dynamischen Modellen. In C. Maurer (Hrsg.), *Naturwissenschaftliche Bildung als Grundlage für berufliche und gesellschaftliche Teilhabe. Gesellschaft für Didaktik der Chemie und Physik, Jahrestagung in Kiel 2018* (S. 464). Universität Regensburg.
- Tesch, M., & Duit, R. (2004). Experimentieren im Physikunterricht—Ergebnisse einer Videostudie. *ZfDN*, 10, 51–69.
- Thees, M., Kapp, S., Strzys, M. P., Beil, F., Lukowicz, P., & Kuhn, J. (2020). Effects of augmented reality on learning and cognitive load in university physics laboratory courses. *Computers in Human Behavior*, 108, 106316.