Pascal Klein[1]
Sebastian Gröber[1]
Jochen Kuhn[1]
Andreas Müller[2]

TU Kaiserslautern[1]
Université de Genève[2]

Videobasierte Aufgaben in den klassischen Übungen zur Experimentalphysik 1- Ergebnisse einer Interventionsstudie

Die Wirksamkeit konventioneller vorlesungsbasierter Physik-Einführungsveranstaltungen wurde in den letzten Jahrzehnten häufig in Frage gestellt und durch zahlreiche, z.T. sehr kostspielige, strukturändernde Innovationen zu verbessern versucht (vgl. Übersicht in Klein et al. 2015a). Das Ziel des Projekts „physics.move“ besteht darin, Instruktionsmaterialien zu entwickeln, die mit geringem organisatorischen Mehraufwand in die bestehende Struktur vorlesungsbegleitender, wöchentlicher Übungen implementiert werden und etablierte traditionelle Aufgabenstellungen um experimentelle Anteile komplementär ergänzen.

Instruktionsmaterial

Verglichen wird (i) eine für die Lehrveranstaltung `universitäre Physik-Übungsgruppe´ *traditionelle* Aufgabenstellung (*T-Aufgaben*) mit (ii) einer darauf basierenden Aufgabe die eine Videoaufzeichnung eines Realexperiments umfasst, welche die Studenten hinsichtlich der jeweiligen physikalischen Fragestellung analysieren und auswerten (*Videoanalyse-Aufgabe, VA-Aufgabe*), sowie (iii) einer Variante in der Studierende darüber hinaus angehalten sind, das Videomaterial unter Nutzen einfacher Freihandexperimente mit geringem Materialaufwand und unter Verwendung mobiler Endgeräte selbst zu erstellen (und natürlich auszuwerten) (*mobile Videoanalyse-Aufgabe, mVA-Aufgabe*). Die konzeptionellen Eigenschaften dieser Aufgabenformate wurden anhand von Beispielen herausgearbeitet und an verschiedenen Stellen publiziert (ebd., Gröber et al. 2014, Klein et al. 2015b), sodass sich dieser Beitrag den bisher gewonnenen empirischen Befunden der im Projekt verorteten Begleitforschung widmet. Abbildung 1 stellt den Verlauf des Projekts schematisch dar.

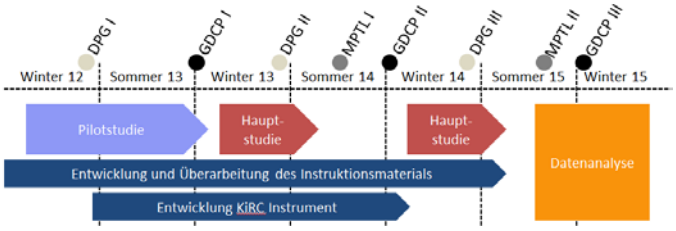

Abb. 1: Schematischer Verlauf einzelner Arbeitsphasen des Projekts physics.move

Bisheriger Forschungsstand im Projekt „physics.move“

In einer ersten Pilotstudie (WiSe 2012/13-SoSe13, siehe Abb.1) wurden sowohl die Implementationsfähigkeit des Ansatzes als auch die Organisation der Feldstudie in einem Kontroll-Versuchsgruppen-Rotationsdesign mit drei Messzeitpunkten belegt (Klein et al. 2015c) sowie Forschungsfragen und Hypothesen für die Hauptuntersuchung formuliert (Klein et al. 2014). Es zeigte sich der Bedarf eines adäquaten Testinstruments zur Erfassung repräsentationaler Kompetenz (RK). Es ist bekannt, dass RK (in einer gegebenen Domäne) ein wesentlicher Faktor für die Entwicklung von Expertise und Problemlösefähigkeit (in dieser Domäne) ist (Meltzer, 2005). Die Übungsaufgaben stellen eine zentrale instruktionale Gelegenheit dar, verschiedene Darstellungsformen zu interpretieren, ineinander überzuführen und selbstständig zu erstellen. Die Überprüfung der Hypothese, dass VA-Aufgaben diese Kompetenz

durch den Einsatz dynamischer Repräsentationen im Besonderen unterstützen (Klein et al. 2014), benötigt ein änderungssensitives Instrument, welches diese Variable operationalisiert. Bisher entwickelte Testinstrumente bilden diese Komponente disziplinärer Kompetenz aber nur unzureichend ab. Wie Nieminen et al. (2010) herausstellen, erlauben bestehende Konzepttests nicht, den Umgang mit verschiedenen Repräsentationsformen hinreichen zu evaluieren. Dadurch konfundiert RK als eine Art „versteckte Variable“ innerhalb bestehender Tests mit der zu messen intendierten Variable, was zu einer ungewünschten Beeinflussung der Testsummen führt. Folglich wurde das KiRC-Instrument zur Erfassung von RK in der Domäne Kinematik entwickelt, überarbeitet, psychometrisch analysiert und validiert (Klein et al. 2016). Neben RK wurden in der Hauptuntersuchung affektive Konstrukte (Neugierde, Motivation) sowie Belastungsvariablen als auch das Konzeptverständnis unter Berücksichtigung mehrerer Kovariaten (Vorleistungen) untersucht. Erste Ergebnisse einer Teilstichprobe (WiSe 2013/14), die positive Befunde in der Entwicklung des akademischen Selbstkonzepts und des Konzeptverständnisses mit mittleren bis hohen Effektstärken im Sinne der Intervention bei unveränderter Belastung attestieren, wurden auf der GDCP-Tagung 2014 präsentiert (Klein et al. 2015b). Zum jetzigen Zeitpunkt ist die Datenerhebung der Hauptstudie (WiSe 2013/14 und WiSe 2014/15) abgeschlossen. In diesem Beitrag präsentieren wir die Ergebnisse der Hypothese zur RK:
Das Arbeiten mit einer Kombination von T-, VA- und mVA-Aufgaben führt zu einem größeren Lernerfolg, d.h. zu einer höheren **Repräsentationskompetenz** (der Kinematik) (RK) als das Arbeiten mit ausschließlich traditionellen Aufgaben.

Kurzbeschreibung des Testinstruments zur Erfassung der Repräsentationskompetenz
Der KiRC-Test beinhaltet 18 Aufgaben zu den grundlegenden Konzepten der Kinematik, die in drei verschiedenen Repräsentationsformen vorliegen (bildhaft, graphisch und algebraisch). Ein Item erfordert beispielsweise die Verknüpfung bildhafter Information (Kugel rollt auf gekrümmter Rampe) mit einem Kinematik-Graphen, während ein anderes Item die algebraisch-formale Interpretation einer eindimensionalen, ungleichmäßig beschleunigten Bewegung abprüft. Mit Hilfe einer Eichstichprobe wurden eine zufriedenstellende Reliabilität (Cronbach's $\alpha = 0.82$), eine substantielle mittlere Trennschärfe (mittlere Item-Testkorrelation $r = 0.28$) und eine ausreichende Diskrimination (mittleres $D = 0.43$) des Gesamtinstruments gemessen, die allesamt kritische Schwellenwerte überschreiten (Ding & Beichner, 2009). Der Test umfasst zwei Itemformate: Multiple-Choice Items (MC3), die neben der richtigen Antwort zwei plausible, aber falsche Alternativen anbieten sowie True-False items (TF), bei denen entschieden werden muss, ob Aussagen richtig oder falsch sind.
Während aus inhaltlicher Sicht eine Kombination beider Formate in einem Test durch Ausnutzen der jeweils spezifischen Vorteile sinnvoll ist, ergibt sich die messtechnische Frage, wie die in den beiden Testhälften erreichten Testsummen T_1 (MC3 Items) und T_2 (TF Items) zu einer gemeinsamen Testsumme T verrechnet werden. Um den verschiedenen Ratewahrscheinlichkeiten sowie der unterschiedlichen Entscheidungsstrategien gerecht zu werden, definieren wir die Testsumme T als

$$\mathrm{T} = \mathrm{T}_1 + \kappa \mathrm{T}_2, \qquad (1)$$

mit dem dimensionslosen Gewichtungsfaktor $\kappa \in [0, \infty)$. Diese Definition erlaubt die Betrachtung einzelner Testhälften ($\kappa = 0$ bzw. $\kappa \to \infty$) und schließt den Spezialfall der heuristischen Gleichgewichtung ein ($\kappa = 1$). Für die Praxis plausibel erscheint der Wertebereich zwischen 0.2 und 1, da TF Items verglichen mit MC Items mit tendenziell niedrigerem, aber nicht zu niedrigem Gewicht eingestuft werden können.

Studiendesign und Stichprobe
Die hier analysierte Kernstichprobe der Hauptuntersuchung (vgl. Abb.1) besteht aus 76 Hauptfach-Studierenden im ersten Physiksemester, zum Prä- und zum ersten Post-Zeitpunkt

an der Befragung teilnahmen und dazwischen vier Aufgabenblätter (mit jeweils vier Übungen) bearbeiteten. Die Kontrollgruppe (KG, N=40) bearbeitete in diesem Zeitraum ausschließlich T-Aufgaben, während die Treatmentgruppe (TG) acht T-Aufgaben, sechs V-Aufgaben und zwei mV-Aufgaben bearbeitete. Die Testung fand unmittelbar vor sowie nach der Intervention statt. Der dritte Messzeitpunkt (nach Gruppenrotation) wird nicht diskutiert.

Ergebnisse

Abbildung 2 (links) zeigt den mittleren Anteil richtiger KiRC-Items für beide Gruppen zu beiden Testzeitpunkten (Gleichgewichtung, $\kappa = 1$). Durch einfaktorielle Varianzanalysen zeigt sich ein signifikanter Zuwachs von RK in beiden Gruppen mit unterschiedlichen Effektgrößen (TG: t (35) = 5.60***, Cohen's d = 0.84, Hake-Faktor g = 0.52; KG: t (39) = 4.26***, d =0.42, g = 0.23). Die erzielten Werte ordnen sich ein in die Ergebnisse über traditionelle Lehre ($g \approx 0.2$) und interaktive konzeptorientierte Methoden (g = 0.3-0.6) ein (Coletta et al. 2007). Der positive Effekt der Intervention zeigt sich auch in signifikanten gruppenspezifischen Unterschieden des Leistungsverlaufes: F (1, 73) = 11.70***, d = 0.80.
Der Interventionserfolg – quantifiziert als Effektstärkemaß – wird in Abb. 2 (rechts) in Abhängigkeit des Gewichtsfaktors κ dargestellt. Über den gesamten Parameterbereich zeichnet sich der Interventionserfolg durch mittlere bis große Effektstärken ab, womit obige Hypothese als bestätigt betrachtet werden kann.

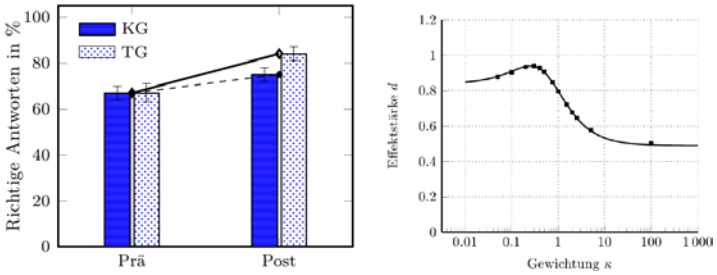

Abb. 2: Ergebnisse bzgl. der Repräsentationskompetenz. Links: Prä/Post-Gruppevergleich für Gleichgewichtung ($\boldsymbol{\kappa = 1}$), rechts: Effektstärke in Abhängigkeit der Gewichtung

Danksagung

Wir danken der Wilfried-und-Ingrid-Kuhn Stiftung für Physikdidaktik für die Förderung des Erstautors und dem Fachbereich Physik der TU KL für die hervorragende Kooperation.

Literatur

David E. Meltzer (2005). Relation between students problem-solving performance and representational format. AJP 73, 463ff

Coletta, V.P., Phillips, J.A. & Steinert, J.J. (2007). Interpreting Force Concept Inventory scores: Normalized gain and SAT scores. Physical Review Special Topics – Physics Education Research, 3(1).

Ding, L. & Beichner, R. (2009). Approaches to data analysis of multiple-choice questions. In: Phys. Rev. ST PER. 5, 020103

Gröber, S., Klein, P. & Kuhn, J. (2014). Video-based problems in introductory mechanics courses. In: EJP 35 055019.

Klein, P., Gröber, S., Kuhn, J. & Müller, A. (2014). Experimentelle Aufgaben in den Übungen zur Experimentalphysik 1. In D. Höttecke (Hrsg.), Naturwissenschaftliche Bildung zwischen Science- und Fachunterricht. GDCP Jahrestagung in München 2013 (S. 312 - 314). Kiel: IPN.

Klein, P., Gröber, S., Kuhn, J. Fouckhardt, H., von Freymann, G., Oesterschulze, E., Widera, A. Fleischhauer, A. & Müller, A. (2015a). physics.move - Videoanalyse-Aufgaben in der Experimentalphysik 1. In: Physik und Didaktik in Schule und Hochschule (PhyDid A) 14 (1)

Klein, P., Gröber, S., Kuhn, J. & Müller, A. (2015b). Experimentelle Aufgaben in den Übungen zur Experimentalphysik 1. In: S. Bernholt (Hrsg.), Heterogenität und Diversität - Vielfalt der Voraussetzungen im naturwissenschaftlichen Unterricht. GDCP Jahrestagung in Bremen 2014 (S. 214-216). Kiel: IPN.

Klein, P., Kuhn, J., Müller, A. & Gröber, S. (2014). Video analysis exercises in regular introductory mechanics courses: Effects of conventional methods and possibilities of mobile devices. In A. Kauertz et al. (Eds.): Multidisciplinary Research on Teaching and Learning, Basingstoke: Palgrave Macmillan

Klein, P., Kuhn, J. & Müller, A. (2016). KiRC inventory: Assessment of representational competence in kinematics. Eingereicht in: Phys. Rev. ST Phys. Educ. Res. (Stand 09/15)

Nieminen, P. Savinainen, A. & Viiri, J. (2010). Force Concept Inventory-based multiple-choice test for investigating students' representational consistency. In: Phys. Rev. ST Phys. Educ. Res. 6, 020109