

Anna B. Bauer
 Peter Reinhold
 Marc D. Sacher

Universität Paderborn

Erhebung der experimentellen Performanz (Physik-)Studierender

Ausgangslage und Ziele

Laborpraktika naturwissenschaftlicher Studiengänge zielen darauf ab, dass Studierende das Experimentieren und insbesondere die wissenschaftliche Erkenntnismethodik erlernen. Dies wird allerdings in den bestehenden Organisations- und Lernumgebungsstrukturen häufig nicht in zufriedenstellendem Maße erreicht (Welzel, 1998).

Vorhandene Testinstrumente (z.B. Arndt, 2016; Rehfeldt & Nordmeier, 2016; Sander, 2000; Straube, 2016) erheben zumeist einzelne Indikatoren erfolgreichen Experimentierens. Im hier vorgestellten Projekt wird dagegen die Tiefenstruktur experimenteller Performanz vor dem Hintergrund des Gesamtprozesses analysiert. Das dabei zu entwickelnde Bewertungsmodell soll neben einer differenzierten Wirkungsanalyse unterschiedlicher Praktikumskonzepte Hinweise für eine Diagnose individueller Fähigkeiten ermöglichen.

Theoretische Grundlagen

Für die Entwicklung des Bewertungsmodells werden Handlungsbeschreibungen generiert, die eine valide Erhebung der studentischen Performanz am Ende des Anfängerpraktikums ermöglichen sollen. Unter Performanz wird in diesem Projekt das nach Neuweg (2011) definierte Können, oder auch *Wissen 3* verstanden. Es wird also aus den gezeigten Handlungen auf die zugrundeliegenden Dispositionen und kognitiven Strukturen geschlossen. Dazu ist es nicht nötig, dass der Kölner sein Können auch verbalisieren kann.

Das wissenschaftliche Experimentieren an Universitäten unterliegt grundsätzlich anderen Zielsetzungen als schulisches Experimentieren (Höttecke, 2015), so dass eine Definition des universitären Experimentierens entwickelt worden ist: Ziel ist das stabilisierende (Galison, 1987) Herauspräparieren der Eigenschaften physikalischer Phänomene (Tetens, 1987), „die in ein kohärentes Verhältnis zu Theorie und Praxis gebracht werden können“ (Höttecke, 2015, S.133). Dafür werden vertieft vernetzte Fähigkeiten und Fertigkeiten zusammen mit höchst flexibel einsetzbaren Wissensbeständen benötigt. Neben dem Entwickeln einer Fragestellung und Hypothesen, dem Planen, Aufbauen, Durchführen und Auswerten experimentell gewonnener Daten (Emden, 2011) nehmen auf universitären Niveau experimentelle Fähigkeiten wie z.B. das Testen und Optimieren experimenteller Aufbauten sowie das differenzierte Beurteilen der Ergebnisse vor dem Hintergrund des Gesamtprozesses einen hohen Stellenwert ein.

Auf universitärer Ebene existieren noch keine differenzierten Fähigkeitsbeschreibungen auf struktureller oder Gesamtprozess-Ebene. Weiterhin fehlt es an Instrumenten zur validen Messung der Qualität der Experimentierhandlungen oder der Differenziertheit der Argumentationen innerhalb des Gesamtprozesses. Heidrich (2017) hat für den Inhaltsbereich Optik versucht Richtigkeit, Zielorientierung und Strukturiertheit als globale Gütekriterien zu definieren und basierend auf videografierten Realexperimenten zu erheben. Dabei konnten für die ersten zwei Gütemaße valide Daten erhoben werden. Für die Strukturierung fehlen noch ausreichende Beschreibungen für mögliche Indikatoren (Heidrich, 2017). An dieser Stelle setzt das hier beschriebene Projekt an. Es werden in einem ersten Schritt anhand der gezeigten Performanz Handlungsbeschreibungen generiert, die eine differenzierte Kategorisierung der Performanz mit Fokus auf die Handlungsumgänge und der damit verbundenen Differenziertheit des experimentellen Vorgehens ermöglichen sollen.

Untersuchungsdesign

Die Entwicklung des Bewertungsmodells wird durch die Erhebung experimenteller Performanz Physikstudierender am Ende des Anfängerlaborpraktikums mit Hilfe videografiert Realexperimente aus dem Inhaltsbereich Elektrodynamik realisiert.

Das Untersuchungsdesign besteht u.a. aus zwei unterschiedlich schweren, teilstrukturierten Aufgabenstellungen. Bei der leichteren Aufgabe haben die Studierenden mit allen Geräten und Inhalten schon experimentell gearbeitet. Mit den Inhalten und den Geräten bei der zweiten, der schwereren Aufgabe ist nur mit einigen Aspekten praktisch oder theoretisch gearbeitet worden. Beide Aufgabenstellungen basieren auf der gleichen Operationalisierung der experimentellen Kompetenz und sind für eine höhere Vergleichbarkeit bis auf die zu beantwortenden Fragestellung identisch strukturiert.

Die Erhebung läuft standardisiert ab: Zunächst werden demografische Daten erhoben. Danach erhalten die Probanden einen Fachwissenstext zur Einführung in ihr Experiment, um den Einfluss des individuellen Fachwissens zu minimieren. Im Anschluss beginnen die Probanden mit dem videografierten Bearbeiten einer der beiden Aufgaben. Sie werden dazu aufgefordert die absolvierten Teilschritte schriftlich zu begründen. Auf Lautes Denkens wurde wegen Validitätsbedenken und zur Senkung der kognitiven Belastung verzichtet (vgl. Konrad, 2014). In der Pilotstudie zeigte Lautes Denken lediglich eine Verbalisierung der Schrittabfolge und nicht der Denkprozesse der Probanden, ist hier also nicht hilfreich. Es wird durch die schriftlichen Begründungen während der Bearbeitung und durch die anschließenden Stimulated Recall Interviews kompensiert. Die Auswertung des Experiments erfolgt an einem Laptop mit dem (aus dem Praktikum bekannten) Programm. An den Laptops kann außerdem im Internet recherchiert werden.

Nach einer ersten Auswertung der Probandendaten findet ein Stimulated-Recall-Interview statt, mit dem die getätigten Rekonstruktionen auf Basis der gezeigten Handlungen abgesichert werden sollen. Hierfür wurde ein standardisierter Interviewleitfaden theoriegeleitet entwickelt. Dabei werden Stimuli aus allen erhobenen Produkten (Video, Aufgabenmitschrift, Auswertungsdatei, Browserverlauf) verwendet.

In der Hauptstudie haben bisher vorrangig Physikstudierende am Ende des Grundpraktikums ($n=8$) sowie ein Doktorand teilgenommen. Im Verlauf des nächsten Semesters werden weitere acht Studierende am Ende des Anfängerpraktikums, Studierende nach Beendigung der Bachelor- und Masterarbeit ($n=2$) sowie Doktoranden ($n=2$) erhoben, um eine möglichst hohe Varianz in den gezeigten Handlungen zu erzielen. Die große Varianz soll die Erstellung der Fähigkeitsniveaus und die Abgrenzung der einzelnen Stufen ermöglichen.

Fallrekonstruktion

Im Folgenden wird beispielhaft aufgezeigt, wie die Auswertung der erhobenen Daten mittels komparativer Fallrekonstruktionen erfolgt. Durch den Ansatz, dass die einzelnen Handlungen vor dem Hintergrund des Gesamtprozesses interpretiert werden, um so Aussagen zu den zugrundeliegenden Dispositionen treffen zu können, ist eine qualitative Untersuchung der Tiefenstruktur erforderlich. Im Fokus stehen somit vor allem die Handlungssequenzen, in denen die Probanden Handlungentscheidungen treffen, da dort die für die Entscheidungsprozesse notwendigen Fähigkeiten und Wissensbestände hinsichtlich ihrer Differenziertheit und Zielorientierung im Gesamtprozess analysiert werden können. Der Ansatz komparativer Fallrekonstruktionen erlaubt es sowohl auf individueller Ebene als auch vergleichend zwischen den Fällen die Tiefenstruktur der gezeigten Performanz zu analysieren. Dafür werden sowohl auf Bestandteile der dokumentarischen Methode (Bohnsack, 2013) für die Rekonstruktion der gezeigten Handlungsmuster (individuell und komparativ) als auch auf Elemente der Objektiven Hermeneutik (Regelgeleitetheit individueller Handlungentscheidungen) zurückgegriffen (Oevermann, 1991).

Im Folgenden wird ein Beispiel für die Rekonstruktion der Fähigkeit (*Mess-)*Geräte

auswählen dargestellt anhand der Daten zweier Probanden (siehe Abb. 1), die zu der Teilaufgabenstellung aus der Phase Planung „Wählen Sie die Messgeräte und Materialien für die Beantwortung der Fragestellung aus und begründen Sie ihre Wahl“ entstanden sind. Proband 1 wählt die (Mess-) Geräte ohne Bezug zu dem zu untersuchenden physikalischen Phänomen aus. Als Kriterium nennt er lediglich deren Bekanntheit. Andere Geräte bezieht er nicht ein. Die Fähigkeit (Mess-)Geräte auszuwählen beherrscht er nicht. Sein Ergebnis beträgt $(8,5 \pm 0,6) \Omega$. Im Interview wird ihm der korrekte Wert von $10 \Omega (\pm 3\%)$ mitgeteilt. Dennoch beurteilt er sein Vorgehen als erfolgreich und würde dieses nicht ändern. Auch der deutliche Mangel hinsichtlich der Präzision des Ergebnisses führt nicht zu einem Überdenken der Messgerätewahl. Er scheint folglich auch die Fähigkeit, Messergebnisse auf Basis von Messgenauigkeiten zu beurteilen, nicht zu beherrschen.

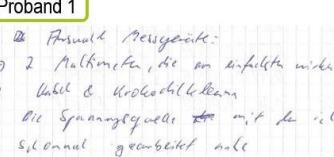
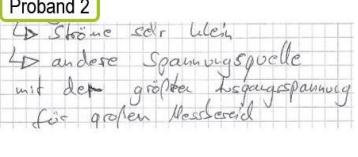
Proband 1  Auszug aus dem Stimulated-Recall-Interview Proband 1: I: Du hast in deinem Laborbuch auf Seite 1 begründet, warum du die Messgeräte gewählt hast. Könntest du mir noch einmal zusammenfassen, was deine Gründe waren? P: Ich habe zwei Multimeter gewählt, die am einfachsten wirkten, die ich schon mal gesehen hatte. I: Ok, das heißt „einfach“ bedeutet in dem Fall nicht simpel sondern „kenne ich schon“? P: Ja, also womit ich am vertrautesten war. I: Hast du dir auch Gedanken gemacht, welche Spannungsquelle du wählst? P: Ne, ich kannte nur die eine. Die andere habe ich mir gar nicht richtig angesehen. [...] I: Würdest du im Nachhinein, deine Messgeräte wieder so wählen, wie du es getan hast? P: Ja, ich habe ja ein Ergebnis herausbekommen.	Proband 2  Auszug aus dem Stimulated-Recall-Interview Proband 2: I: Du hast in deinem Laborbuch auf Seite 3 begründet, warum du die Messgeräte gewählt hast. Könntest du mir zusammenfassen, was deine Gründe waren? P: Ich habe zuerst die Größenordnung der Widerstände bestimmt [mit einem Multimeter] und mir dann ja überlegt, welche Messbereiche und Intervalle ich mit den Multimetern messen muss und so. Ja und dann habe ich mich erinnert, dass Widerstände [Bauteil] eine Leistungsgrenze haben. Und dann habe ich das durchgerechnet und festgelegt bis wohn ich messen muss. Ja, ne und dann habe ich die zwei Multimeter genommen. I: Ok, ich habe dann im Video gesehen, dass du auch zwei verschiedene Spannungsquellen verwendet hast. P: Ja, ich wusste, dass der zweite Widerstand im MQ-Bereich liegt und da habe ich die andere Spannungsquelle gewählt, weil ich dann einen größeren Messbereich habe. Ja. Finde ich auch immer noch logisch.
---	---

Abb. 1: Beispielhafte Auszüge aus der Aufgabenmitschrift und dem Stimulated Recall Interview zweier Probanden.

Proband 2 wählt die Messgeräte mit Bezug zum physikalischen Phänomen aus, führt eine Testmessung durch und prüft die Eignung aller Geräte. Als Auswahlkriterien führt er neben Eigenschaften des Widerstandes (Bauteil) auch die Messbereiche der Multimeter an. Die Fähigkeit Messgeräte auszuwählen beherrscht er elaboriert. Als Ergebnis bestimmt der Student den Wert $(9,81 \pm 0,04) \Omega$. Nach Mitteilung des korrekten Wertes (nicht in Abb. 1 enthalten) beurteilt er sein Vorgehen auf Basis der bestimmten Messunsicherheit als erfolgreich und würde dieses nicht ändern. Er scheint folglich auch die Fähigkeit Messergebnisse auf Basis von Messunsicherheiten zu beurteilen zu beherrschen.

Im Vergleich verfügt Proband 2 über elaboriertere Fähigkeiten als Proband 1. Er zeigt eine hohe Differenziertheit bei seinen Analysen und Beurteilungen seiner Handlungsschritte sowie der erhobenen Daten. Proband 1 hingegen beschränkt sich auf seine bisherigen Erfahrungen und zeigt kaum differenzierte Handlungsentscheidungen. Er hat den Gesamtprozess des Experimentierens nicht im Blick, stattdessen arbeitet er lediglich die Aufgabenstellung ab und ist zufrieden ein Ergebnis erreicht zu haben.

Ausblick

Im Laufe des nächsten Jahres soll die Hauptstudie abgeschlossen werden. In einem weiteren Schritt sollen mit Hilfe einer typenbildenden Inhaltsanalyse (Kuckartz, 2012) Niveaustufen gebildet werden, welche mit Hilfe eines Expertenratings abgesichert werden sollen.

Literatur

- Arndt, K. (2016). Experimentierkompetenz erfassen. Analyse von Prozessen und Mustern am Beispiel von Lehramtsstudierenden der Chemie: Dissertation. Berlin: Logos Berlin (Studien zum Physik- und Chemielernen, 202).
- Bohnsack, R.; Nentwig-Gesemann, I.; Nohl, A. (Hg.) (2013). Die dokumentarische Methode und ihre Forschungspraxis. Grundlagen qualitativer Sozialforschung. 3., aktualisierte Aufl. Wiesbaden: Springer VS.
- Emden, Markus (2011). Prozessorientierte Leistungsmessung des naturwissenschaftlich-experimentellen Arbeitens. Eine vergleichende Studie zu Diagnoseinstrumenten zu Beginn der Sekundarstufe I. Berlin: Logos Berlin (Studien zum Physik- und Chemielernen, 118).
- Galison, P. L. (1987). How experiments end. Chicago: University of Chicago Press.
- Heidrich, J. (2017). Erfassung von Experimentierkompetenz im universitären Kontext. Entwicklung und Validierung eines Experimentiertests zum Themenbereich Optik. Kiel. Online verf.: http://macau.uni-kiel.de/servlets/MCRFileNodeServlet/dissertation_derivate_00007080/DissHeidrich.pdf.
- Höttinge, D.; Rieß, F. (2015). Naturwissenschaftliches Experimentieren im Lichte der jüngeren Wissenschaftsforschung – Auf der Suche nach einem authentischen Experimentbegriff der Fachdidaktik. Zeitschrift für Didaktik der Naturwissenschaften. In: ZfDN 21 (1), 127–139.
- Konrad, K. (2010). Lautes Denken. In: Günter Mey und Katja Mruck (Hg.): Handbuch Qualitative Forschung in der Psychologie. Wiesbaden: VS Verlag für Sozialwissenschaften / Springer Fachmedien Wiesbaden GmbH Wiesbaden, 476–490.
- Kuckartz, U. (2012). Qualitative Inhaltsanalyse Methoden, Praxis, Computerunterstützung (2. Aufl.). Weinheim: Beltz Juventa.
- Neuweg, G. H. (2011). Das Wissen der Wissensvermittler. In: Terhart, E., Bennewitz, H., Rothland, M. (Hg.): Handbuch der Forschung zum Lehrerberuf. Münster: Waxmann, 451–477.
- Oevermann, U. (1991). Genetischer Strukturalismus und das sozialwissenschaftliche Problem der Erklärung der Entstehung des Neuen. In: S. Müller-Dohm (Hg.), Jenseits von Utopia. Theoriekritik und Gegenwart. Frankfurt/M., 267–336.
- Rehfeldt, D.; Nordmeier, V. (2016). Skalen zur Messung von Praktikumsqualität: Konfirmatorische Analyse der Struktur und Konstrukte. Didaktik der Physik: Frühjahrstagung, Hannover.
- Sander, F. (2000). Verbindung von Theorie und Experiment im physikalischen Praktikum. Eine empirische Untersuchung zum handlungsbezogenen Vorverständnis und dem Einsatz grafikorientierter Modellbildung im Praktikum. Univ., Diss.-Bremen, 1999. Berlin: Logos-Verl. (Studien zum Physiklernen, 13).
- Straube, P. (2016). Modellierung und Erfassung von Kompetenzen naturwissenschaftlicher Erkenntnisgewinnung bei (Lehramts-) Studierenden im Fach Physik. Dissertation. Logos Verlag Berlin.
- Tetens, H. (1987). Experimentelle Erfahrung. Eine wissenschaftstheoretische Studie über die Rolle des Experiments in der Begriffs- und Theoriebildung der Physik. Teilw. zugl.: Marburg, Univ., Habil.-Schr., 1986. Hamburg: Meiner (Paradeigmata, 8).
- Weinert, F. E. (2001). Concept of competence: A conceptual clarification. In D. S. Rychen & L. H. Salganik (Eds.), Defining and selecting key competencies, 45–65. Ashland, OH, US: Hogrefe & Huber Publishers.
- Welzel, M.; Haller, K. et al. (1998). Ziele, die Lehrende mit dem Experimentieren in der naturwissenschaftlichen Ausbildung verbinden. Ergebnisse einer europäischen Umfrage. In: Zeitschrift für Didaktik der Naturwissenschaft 4 (1), 29–44.