

Daniel Rehfeldt
Tobias Gutzler
Volkhard Nordmeier

Freie Universität Berlin

TSL: Quantitative Problemanalyse im Nebenfachpraktikum

Das Projekt Technology SUPPORTed Labs, kurz TSL, hat die Verbesserung naturwissenschaftlicher Experimentalpraktika unter Multimediaeinsatz an der Freien Universität Berlin zum Ziel – unter anderem mit Fokus auf das physikalische Praktikum für Naturwissenschaftler (NP), das Studierende im zweiten Fachsemester im Nebenfach belegen. Es ist organisatorisch als klassisches „Anfängerpraktikum“ zu bezeichnen (Rehfeldt, Gutzler, & Nordmeier, 2013). Als Forschungsparadigma für die vorliegende quantitative Studie zur Problemanalyse im NP wurde der Design-Based-Research-Ansatz, kurz DBR, nach Reinmann (2005) gewählt, da er Vorteile der Evaluationsforschung mit Theoriebildung kombiniert und speziell für multimediale Interventionen, wie im vorliegenden Fall, geeignet ist. Der erste Schritt beim DBR ist die Problemanalyse.

Probleme von Nebenfachpraktika

Es liegen einige umfassende Forschungsergebnisse zu den physikalischen Praktika für Mediziner vor (Plomer, 2011; Theyßen, 2000). Allerdings mangelt es an einer umfassenden Problemanalyse in anderen Nebenfachpraktika (Chemie, Biochemie, Geowissenschaften etc.), in der nicht á priori Problembereiche durch den Forschenden festgelegt sind oder durch subjektive Erfahrungen desselben festgelegt werden.

Innerhalb dieses Praktikumstyps wurden zudem noch nie Lehrende und Lernende zugleich nach denselben Problemen befragt. Dies ist nach Gollwitzer & Jäger (2009, S. 141ff.) jedoch für eine objektive Erfassung der Probleme gewinnbringend, da eine Aggregation von unterschiedlichen Einschätzungen als valider Indikator für latente Konstrukte dienen kann.

Forschungsfrage und Hypothesen

Die Forschungsfrage bezieht sich demzufolge auf eine Gesamtanalyse der Probleme des physikalischen Praktikums für Naturwissenschaftler: *Welche Probleme des NP sind sowohl für die TeilnehmerInnen, als auch für die BetreuerInnen schwerwiegend?* Aus der qualitativen Vorerhebung (Gutzler, Rehfeldt, & Nordmeier, in diesem Tagungsband) wurden acht zentrale Befunde extrahiert. Diese stellen im Rahmen von TSL Evaluationsgegenstände dar und erfordern ggf. Interventionen mit didaktischem Fokus. Zugleich sind dies die zentralen Forschungshypothesen H1–H8 der vorliegenden Studie, die in Abbildung 1 dargestellt sind.

physikalisch-methodisches Vorwissen (4 Items, $\alpha = .83$) Fragehäufigkeit zur Theorie unzureichendes physikalisches Vorwissen (H1)	Skriptqualität: Vorbereitung auf Versuch (2 Items, $\alpha = .87$) Platzskripte: Anleitung Versuch unzureichende Vorbereitung auf praktische Versuchsteile (H3)	Kurzttest: Motivation für Beschäftigung mit Aufbau Fragehäufigkeit zu Experimentiergeräten gezielte Vorbereitung auf experimentelle Aufgaben	Fragenbekanntheit Kurzttest Kurzttest: Motivation für Theorie Rückmeldungswunsch bzgl. Kurzttest (2 Items, $\alpha = .73$)
Experimentier- & Gerätekenntnisse (3 Items, $\alpha = .83$) unzureichende Experimentier- & Gerätekenntnisse (H2)	Arbeitsbelastung NP (2 Items, $\alpha = .83$) Erfolgsdruck NP (2 Items, $\alpha = .87$) insgesamt zu hohe Arbeitsbelastung im NP (H6)	protokollbezogene PC-Affinität (3 Items, $\alpha = .75$) durchschnittliche Protokollnote qualitativ schlechte PC-Protokolle (H8)	Kurzttest: Motivierend für Theorie & Überarbeitungsbedürftig (H4) Betreuungsqualität bzgl. Durchführung (2 Items, $\alpha = .72$) unzureichende Betreuung während des Versuchs (H7)
Erfolgsdruck Kurzttest (2 Items, $\alpha = .87$) Kurzttest: Erhöhter Erfolgsdruck (H5)			

Abb. 1: Forschungshypothesen mit Skalen und Itembezeichnungen für die Vollerhebung

Konzeption und Validierung des Testinstruments

Die Problemanalyse wurde multimethodisch (Gollwitzer & Jäger, 2009, S. 141ff.) als Vollerhebung (Stichprobe bestehend aus TeilnehmerInnen und BetreuerInnen) durchgeführt. Als inhaltliche Grundlage für die Fragebogenkonstruktion¹ dienten die zentralen Befunde aus der qualitativen Vorerhebung. Die Konstrukte für den Fragebogen setzten sich aus etablierten Skalen (z. B. Arbeitsbelastung) und faktoriell validierten Skalen (z. B. Experimentierkenntnis) zusammen (vgl. Abb. 1). Für die Erfassung der Teilnehmenden- und Betreuendenperspektive wurde mit parallel formulierten Items gearbeitet. Verständlicherweise gibt es einige Problembereiche, die ausschließlich die TeilnehmerInnen valide bewerten können, wie etwa die Wahrnehmung von Erfolgsdruck. Für diese Probleme wurden teilweise Kontrollitems bei den Betreuenden eingesetzt.

Zur Validierung des Testinstruments wurde mit den Betreuenden und Teilnehmenden des Praktikums ein Zwei-Phasen-Pretest (kognitive Interviews & empirischer Pretest) durchgeführt. Mit der Think-Aloud-Technik wurden die Itemformulierungen in ihrer Verständlichkeit zielgruppenspezifisch verbessert, mit dem empirischen Pretest die induktiv konstruierten Skalen validiert (Itemanalyse und explorative Faktorenanalyse).

Erhebung

Die Befragung wurde innerhalb der Praktikumszeit auf mobilen Endgeräten durchgeführt. Insgesamt konnten so 121 von 131 TeilnehmerInnen und 18 von 23 BetreuerInnen für die Erhebung gewonnen werden. Die TeilnehmerInnen befanden sich in der Regel im zweiten Fachsemester (Bachelor) in einem der Studiengänge Chemie, Biochemie oder Geowissenschaften. Die BetreuerInnen waren allesamt Studierende der Physik bzw. Physik Lehramt.

Auswertungsverfahren

Ein Problem wird dann als „schwerwiegend“ gemäß Forschungsfrage definiert, wenn bei einer Problemskala (vgl. Abb. 2) eine statistisch signifikante Abweichung vom Cutoffwert 3.5 hypothesenkonform gerichtet vorliegt.² Der Cutoffwert repräsentiert ein als mittelschwer wahrgenommenes Problem, da dem Problemvorliegen weder klar zugestimmt, noch es komplett abgelehnt wird. Eine Bestätigung der Hypothese liegt dann vor, wenn die Testung signifikant und mit mindestens mittlerer Effektstärke ($\bar{r} \geq .3$, Bortz, 2010, S. 165) in mindestens der Hälfte der relevanten Skalen/Items verläuft und sich keine hypotheseninkonformen Effekte zeigen.



Abb. 2: Beispiel-Item mit Zuordnung des Cutoff-Wertes 3.5

Ergebnisse

Es wurden drei von acht Problemen klar bestätigt (vgl. Tab. 1). Bei den signifikant von 3.5 abweichenden Skalenmittelwerten lag je ein Signifikanzniveau von $p < .001$ vor. Der Begriff „tendenziell“ kennzeichnet Skalen, die sich unter der BONFERRONI-Korrektur nicht von 3.5 unterscheiden.

Für die Hypothese H3 zeigt die Auswertung eine mittlere Effektstärke, H4 ergibt aus BetreuerInnensicht sogar starke Effekte. H6 konnte inhaltlich nur aus Studierendensicht bewertet werden und ist daher etwas konservativer zu interpretieren.

¹ rational-induktive Testkonstruktion, Bühner, 2006, S. 47ff

² 1-Stichproben-t-Test mit BONFERRONI-Korrektur, Bortz, 2010, S. 118

Tabelle 1 zeigt genauer, aus welcher Perspektive die Probleme bestätigt wurden.

Tab. 1: Ergebnisse der Problemanalyse

Problem	schwerwiegend aus TN-Perspektive?	\bar{r}	schwerwiegend aus BETR-Perspektive?	\bar{r}
H1: geringes Vorwissen	Nein	.	Ja	.
H2: geringe Gerätekenntnis	Ja	.	Nein	.
H3: schlechte Vorbereitung auf Praxisteil insgesamt	Ja: 3 von 4 Skalen	.28	Ja: 2 von 4 Skalen	.38
H4: Kurztest: motivierend, aber überarbeitungsbedürftig	Ja: 2 von 3 Skalen	.47	Ja: 2 von 2 Skalen	.74
H5: Kurztest: hoher Erfolgsdruck	Ja: tendenziell	.	nicht erfasst	.
H6: hohe Arbeitsbelastung & Erfolgsdruck insgesamt	Ja: 2 von 2 Skalen	.69	nicht erfasst	.
H7: schlechte Betreuung	Nein	.	Nein (Kontrollitems)	.
H8: schlechte PC-Protokolle	Nein	.	Nein (Kontrollitems)	.

Legende: TN – TeilnehmerInnen, BETR – BetreuerInnen

Fazit & Ausblick

Mit Hilfe der quantitativen Problemanalyse des NP konnten (gemäß den angelegten Kriterien) als Antwort auf die eingangs gestellte Forschungsfrage drei als „schwerwiegend“ zu bezeichnende Problemfelder bestätigt bzw. identifiziert werden:

1. Die Vorbereitung bereitet zu wenig auf den praktischen Teil des Versuchs vor.
2. Der Kurztest ist motivierend für die Erarbeitung der Theorie, aber überarbeitungsbedürftig.
3. Die Arbeitsbelastung im NP ist zu hoch.

Im Rahmen einer nun folgenden Intervention soll daher vor allem auf eine Stärkung der Vorbereitung hinsichtlich praktischer experimenteller Aspekte abgezielt werden (z. B. durch den Einsatz von IBE). Der Kurztest kann für die Theorieerarbeitung erhalten bleiben, allerdings ist eine Generalüberholung der Testfragen unumgänglich. Hinsichtlich des Problems „Arbeitsbelastung und Erfolgsdruck“ sollte das NP auch auf organisatorischer Ebene überdacht werden. Auf diesen empirischen Befunden aufbauend folgt in einem nächsten Schritt die Konzeption von Maßnahmen entlang der drei schwerwiegenden Problemfelder. Gemäß DBR wird dies in einem iterativen Prozess in Zusammenarbeit mit Beteiligten des Praktikums erfolgen.

Das Projekt TSL wird im Rahmen des Programms SUPPORT der Freien Universität Berlin durch das BMBF gefördert.

Literatur

- Bortz, J. (2010). Statistik für Human- und Sozialwissenschaftler (7., vollständig überarbeitete und erweiterte Auflage.). Berlin: Springer.
- Bühner, M. (2006). Einführung in die Test- und Fragebogenkonstruktion (2., aktualisierte Auflage.). Pearson.
- Gollwitzer, M., & Jäger, R. S. (2009). Evaluation kompakt (1. Aufl.). Beltz.
- Plomer, M. (2011). Physik physiologisch passend Praktiziert. Berlin: Logos.
- Rehfeldt, D., Gutzler, T., & Nordmeier, V. (2013). TSL – Technology SUPPORTed Labs: Multimediale Unterstützung naturwissenschaftlicher Hochschulpraktika. In V. Nordmeier & H. Grötzebauch (Hrsg.), PhyDid B, Didaktik der Physik, Beiträge zur DPG-Frühjahrstagung, Jena.
- Reinmann, G. (2005). Innovation ohne Forschung? Ein Plädoyer für den Design-Based Research-Ansatz in der Lehr-Lernforschung. Unterrichtswissenschaft, 33(1), 52–69.
- Theyßen, H. (2000). Ein Physikpraktikum für Studierende der Medizin. Berlin: Logos.