

Der Einfluss universitär vermittelter professioneller Kompetenzen auf das Handeln in unterrichtlichen Erklärsituationen

Der Zusammenhang von Professionswissen bzw. professioneller Kompetenz von Physiklehrkräften einerseits und der Handlungsqualität in Unterrichtssituationen ist empirisch nicht geklärt (Vogelsang & Reinhold, 2013; Cauet et al., 2015). Das ist ein Kernproblem für die Lehrerbildung. Die gesamte (universitäre) Ausbildung hat das Ziel, professionelles Wissen bzw. professionelle Kompetenzen in den drei Bereichen Fach, Fachdidaktik und Pädagogik zu entwickeln. Quasi-Längsschnitte legen auch nahe, dass im Verlaufe des Studiums solche Entwicklungen stattfinden (z.B. Riese, 2009). Ob diese drei Bereiche jedoch überhaupt Ressourcen sind, aus denen eine Physiklehrkraft während des unterrichtlichen Handelns gewinnen kann, ist nicht genügend geklärt – noch weniger klar ist die Frage, welche Aspekte professionellen Wissens für welche Handlungssituation verwendet werden können. Als Folge ist so auch unbekannt, ob und unter welchen Bedingungen Quereinsteigerprogramme zu einem Standardweg werden können, den Lehrberuf zu ergreifen bzw. in welchen Bereichen Quereingestiegende nachgeschult werden müssen. Ebenfalls stellt sich die Frage, welche Inhalte empirische gestützt im Lehramtsstudium verankert werden müssen; bislang werden diese Inhalte nahezu ausschließlich normativ und tradiert festgelegt. Ein Grund dafür, dass der Nachweis schwierig ist, liegt auch in der Art und Weise, wie üblicherweise getestet wird.

Testformate für professionelles Wissen, Kompetenz und Unterrichtshandeln

Der gängige Weg, professionelles Wissen bzw. professionelle Handlungskompetenz zu erheben, ist ein schriftlicher Test. In Anlehnung an Miller (1990) können vier Formate des Testens unterschieden werden, wenn berufliche Kompetenzen erhoben werden sollen. Sie gestalten sich bezüglich des Lehrberufs mit unterschiedlichen Vor- und Nachteilen:

1. *Tests des Wissens.* Hierbei handelt es sich in der Regel um schriftliche Testformate mit geschlossenen Aufgaben. Sie sind zumeist mit großem Fokus auf curriculare Validität entwickelt worden.
2. *Tests der Kompetenz.* Auch dies sind üblicherweise schriftliche Testformate, oft mit Unterrichtsvignetten, die die Anwendung von Wissen auf Probleme des Physikunterrichts erfordern. Auch hier ist der Fokus vor allem curriculare Validität, d.h. die normativen Inhalte des Lehramtsstudiums.
3. *Tests des freien beruflichen Handelns.* Dies sind in der Regel Videographien von Unterrichtsstunden. Unterricht besteht aus einer Vielzahl an hochkomplexen Problemen, die unter großem Handlungsdruck gelöst werden müssen. Die Auswertung ist aufwändig und es gelingt auch in großen Studien nicht, eine große Anzahl an Stunden pro Lehrkraft zu filmen.
4. *Tests der Performanz.* Hierbei handelt es sich auch um Videobeobachtungen, allerdings werden hierbei Lehrkräfte mit standardisierten Problemen unter standardisierten Rahmenbedingungen konfrontiert.

Wissens- und Kompetenztests lassen sich mit hoher Testökonomie und hoher Reliabilität konstruieren. Sie lassen sich ebenfalls so entwickeln, dass sie den Lernzuwachs im Verlaufe der Lehrerbildung abbilden. Der Zusammenhang zu mit Unterrichtsvideos erhobenen Daten zur Unterrichtsqualität ist jedoch nicht gelungen. Ein Grund dafür ist, dass Unterrichtsbeobachtungen immer einer Vielzahl an nichtkontrollierbaren Rahmenbedingungen unterworfen sind: So hängt das Gelingen einer Stunde beispielsweise

von der Tagesform der Schülerinnen und Schüler, der Lage einer Stunde oder auch dem Klassenklima ab. Werden all diese Rahmenbedingungen als zufällig schwankend angenommen, so bedarf es einer großen Anzahl an Stunden, die gefilmt werden, bis der „echte“ Effekt beispielsweise des fachdidaktischen Wissens gezeigt werden kann, da sich dann Fehler durch die Rahmenbedingungen erwartbar herausmitteln. Wegen des großen Aufwands ist dies jedoch selbst sehr groß angelegten Projekten nicht möglich. Performanztests hingegen kontrollieren die Rahmenbedingungen – ihr großer Nachteil ist allerdings, dass sie kein holistisches Bild vom Handeln zeigen, sondern nur die Handlungsqualität auf eine bestimmte Unterrichtssituation bezogen. Im Gegensatz zum freien beruflichen Handeln konzentrieren sie sich auf ein einziges berufliches Problem, z.B. das Erklären von Physik. Um einem ganzheitlichen Bild unterrichtlicher Fähigkeiten nahezukommen, müssten alle Standardsituationen des Physikunterrichts in Performanztests abgebildet werden – das Projekt ProfiLe-P+ (siehe Vogelsang in diesem Band) arbeitet daran. In der Medizin sind Performanztests verbreitet, z.B. wird die Aufnahme einer Krankengeschichte simuliert oder die Untersuchung des Bauchraums. Sie werden unter dem Schlagwort „Objective structured clinical examination“ zusammengefasst (Harden et al., 1975). Ihr Kern sind Personen, die mit festen Rollenbeschreibungen Patienten darstellen („standardized patients“) (Barrows & Abrahamson, 1964). Es lässt sich auch zeigen, dass mit Score Sheets in diesen Formaten mit hoher Testgüte Leistungen festgestellt werden können (Walters, Osborn & Raven, 2005).

Ein Performanztest für Erklären im Physikunterricht

Im Projekt ProfiLe-P wurde ein Performanztest entwickelt, der eine bestimmte Unterrichtssituation nachstellt: das Erklären von Physik. Analog zu den „standardized patients“ in medizinischen Performanztests ist der Kern dieses Tests ein standardisierter Schüler, dessen Verhalten einer Rollenbeschreibung folgt: zum Beispiel werden in allen Tests bestimmte Fragen gestellt, entweder, um Nichtverstehen zu äußern (z.B. „Gibt es dafür auch ein Beispiel?“) oder um die Erklärung komplexer zu gestalten (z.B. „Gibt’s da auch eine Formel für? Ich brauche das für eine Klausur.“). Dadurch werden die Tests miteinander vergleichbar. Die Schüler wurden mit Videofeedbackverfahren in etwa zwei bis drei Stunden dazu trainiert, sich vergleichbar zu verhalten. Ebenfalls standardisiert ist das Thema (z.B. „Warum gleitet man auf einer Pfütze eher aus einer Kurve als auf trockener Strecke?“), die Zeit zur Vorbereitung (10 min) sowie die Erklärdauer (10 min). Auch Materialien wie Diagramme und Zeichnungen stehen in allen Szenarien gleichermaßen bereit.

Der Test wurde breit angelegten Validierungsstudien unterzogen, z.B. wurden Interviewstudien durchgeführt, um zu kontrollieren, ob die Situationen authentisch sind. Zur Auswertung der Handlungsqualität wird auf ein Modell des Erklärens von Physik zurückgegriffen (Kulgemeyer & Schecker, 2012). Dabei werden die Videos kategorienbasiert ausgewertet und die positiv zur Erklärqualität beitragenden Kategorien summiert (z.B. verwendete Beispiele). Das so entstehende Maß wurde mit Expertenbefragungen und der Analyse eines nomologischen Netzwerks weiteren Validierungsstudien unterzogen. Im Detail ist die Testentwicklung dargestellt in Kulgemeyer und Tomczyszyn (2015).

Der Zusammenhang von universitär erworbenen Kompetenzen und Handlungsqualität

Dieser Test wurde im Rahmen des Projekts ProfiLe-P verwendet, um den Einfluss von universitär erworbenem Wissen bzw. Kompetenzen im Bereich Fach und Fachdidaktik auf das Handeln beim Erklären zu untersuchen. Die Kernfrage dabei ist, ob Fachwissen und fachdidaktisches Wissen genutzt werden können, um die Erklärqualität zu steigern. Dazu wurden insgesamt 198 angehende Physiklehrkräfte aller Semester an fünf Universitäten befragt. 134 davon haben neben dem Performanztest zum Erklären Tests zum Fachwissen

(entwickelt an den Universitäten Duisburg-Essen und Potsdam) und zum fachdidaktischen Wissen (entwickelt an der Universität Paderborn) ausgefüllt, zudem wurden einige Kontrollvariablen erhoben (z.B. mathematische Fähigkeiten, epistemologische Überzeugungen, Selbstkonzept, etc.). Einen Überblick über die Anlage der Studie und die Validierungsstudien findet sich in Riese et al. (2015).

Es ergeben sich zunächst manifeste Korrelationen zwischen der Erklärperformanz und sowohl Fachwissen ($r = 0,376$; $p < 0,01$) als auch fachdidaktischem Wissen ($r = 0,376$; $p < 0,001$). Zur genaueren Analyse des Einflusses wurde ein manifestes Pfadmodell berechnet. Dabei wurde insbesondere überprüft, ob ein direkter Effekt vom Fachwissen zur Erklärperformanz vorliegt oder ein indirekter Effekt über das fachdidaktische Wissen. Das Modell, das die besten Fitwerte zeigt, ist in Abb. 1 angegeben und erklärt insgesamt 29 % der Varianz in der Erklärperformanz. Dabei ist bemerkenswert, dass Fachwissen nicht direkt auf die Erklärperformanz wirkt, sondern durch fachdidaktisches Wissen mediiert wird.

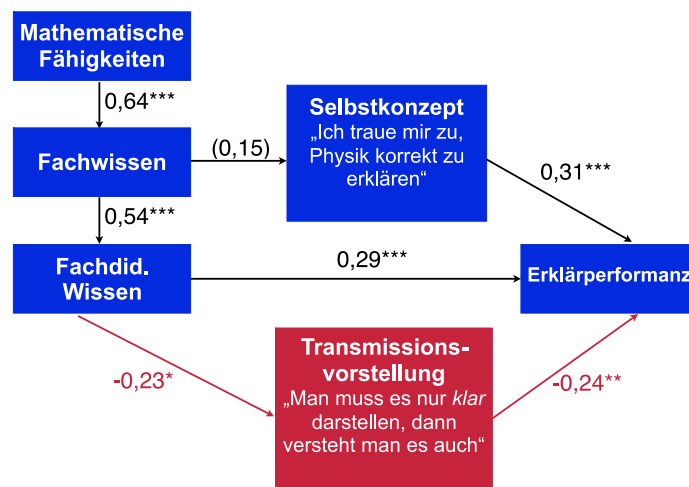


Abb. 1: Pfadmodell (Robust Maximum Likelihood Estimation). $\chi^2/df = 1,1$ ($p = 0,363$), $CFI = 0,994$, $RMSEA = 0,03$; $R^2 = 29\%$. Die Zahlen unter den Pfeilen entsprechen Pfadkoeffizienten (* = $p < 0,05$; ** = $p < 0,01$; *** = $p < 0,001$).

Ebenfalls beachtenswert ist der starke Einfluss von Einstellungen auf die Handlungsqualität. Das auf Erklärsituationen bezogene Selbstkonzept hat einen robusten, positiven Einfluss. Einen negativen Einfluss hat die Transmissionsvorstellung als Aspekt epistemologischer Überzeugungen: Personen, die Erklären als „klares Darstellen“ und nicht als am Adressaten orientierte Handlung auffassen, erklären schlechter. Bei der weiteren Analyse zeigt sich, dass diese Personen oft auch ein hohes Interesse am Erklären haben. Das lässt sich so interpretieren, dass Erklären oft als dozierende Handlung missverstanden wird und nicht als Interaktion zwischen Erklärer und Adressaten.

Mit diesem Performanztest konnte gezeigt werden, dass Fachwissen und fachdidaktisches Wissen einen positiven Einfluss auf das Handeln in einer unterrichtlichen Situation, nämlich dem Erklären, haben. Da sowohl Fachwissen als auch fachdidaktisches Wissen curricular valide erhoben wurde, liegt es nahe, dass die im Lehramtsstudium vermittelten Inhalte zumindest für Erklärsituationen wertvoll genutzt werden können. Insbesondere ist fachdidaktisches Wissen notwendig, solitäres Fachwissen genügt explizit nicht. Dies ist auch für Quereinsteigerprogramme eine relevante Erkenntnis. Weitere korrelative Analysen zeigen, dass insbesondere das Wissen über Schülervorstellungen im Bereich des fachdidaktischen Wissens besonders wichtig ist.

Literatur

- Barrows, H. & Abrahamson, S. (1964). The Programmed Patient: A Technique for Appraising Student Performance in Clinical Neurology. *Journal of Medical Education* 39(8), S. 802–805.
- Cauet, E., Liepertz, S., Borowski, A. & Fischer, H. (2015). Does it matter what we measure? Domain-specific professional knowledge of physics teachers. *Revue suisse des sciences de l'éducation* 37(3), S. 462–479.
- Harden, R., Stevenson, M. & Wilson, W. (1975). Assessment of Clinical Competence using Objective Structured Examination. *British Medical Journal* 1, S. 447–451.
- Kulgemeyer, C. & Schecker, H. (2012). Physikalische Kommunikationskompetenz – Empirische Validierung eines normativen Modells. *Zeitschrift für Didaktik der Naturwissenschaften* 18, S. 29–54.
- Kulgemeyer, C. & Tomczyszyn, E. (2015). Physik erklären – Messung der Erklärfähigkeit angehender Physiklehrkräfte in einer simulierten Unterrichtssituation. *Zeitschrift für Didaktik der Naturwissenschaften* 21(1), S. 111–126.
- Miller, G. (1990). The Assessment of Clinical Skills/Competence/Performance. *Academic Medicine* 65(9), S. 563–567.
- Riese, J. (2009). *Professionelles Wissen und professionelle Handlungskompetenz von (angehenden) Physiklehrkräften*. Berlin: Logos.
- Riese, J., Kulgemeyer, C., Borowski, A., Fischer, H., Gigl, F., Gramzow, Y., Schecker, H., Tomczyszyn, E. & Zander, S. (2015). Modellierung und Messung des Professionswissens in der Lehramtsausbildung Physik. *Zeitschrift für Pädagogik* (61. Beiheft), S. 55–79.
- Vogelsang, C. & Reinhold, P. (2013). Zur Handlungsvalidität von Tests zum professionellen Wissen von Lehrkräften. *Zeitschrift für Didaktik der Naturwissenschaften* 19, S. 129–157.
- Walters, K., Osborn, D. & Raven, P. (2005). The development, validity and reliability of a multimodality objective structured clinical examination in psychiatry. *Medical Education* 39, S. 292–298.