

Objective Structured Teaching Examinations (OSTE) im Lehramtsstudium

Einleitung

Prüfungen im Rahmen des Lehramtsstudiums können - je nach konkreter Implementation - das Lernverhalten Studierender maßgeblich beeinflussen (Flores, Veiga Simão, Barros & Pereira, 2015). Sie können Studierende durch Rückmeldungen über ihren Leistungsstand und Lernerfolg informieren und so beim Erreichen von zukünftigen Lernzielen unterstützen (Nicol & Macfarlane-Dick, 2006; Sopka, Simon & Beckers, 2018). Prüfungen können so auch als Lerngelegenheiten fungieren („*assessment for learning*“, Carless, Bridges, Chan & Glofcheski, 2017), die gemeinsam mit Lernzielen und Lernaktivitäten einen kohärenten Lernprozess bilden (s. *Constructive Alignment*, Biggs & Tang, 2011) und dazu beitragen theoretisches Wissen und praktische Handlungsanforderungen miteinander zu verbinden (vgl. Zeichner, 2010). Dies geschieht jedoch im Lehramtsstudium durch oft handlungsferne Prüfungsformate nur selten. Während der ersten Ausbildungsphase (Studium) dominieren Prüfungsformate, die deklaratives (*knows*) und prozedurales Wissen (*knows how*, s. *Framework for clinical assessment*, Miller, 1990) fokussieren, wie ein Großteil der Klausuren und mündlichen Prüfungen. In der zweiten Ausbildungsphase (Vorbereitungsdienst) hingegen wird im Rahmen von Unterrichtsbesuchen und später der unterrichtspraktischen Prüfung das Handeln in realen beruflichen Situationen geprüft (*does*). Hier besteht also eine Art „Prüfungslücke“ zwischen der Prüfung von Wissen einerseits und beruflichem Handeln in realen Situationen andererseits. Miller (1990) bezeichnet diese Lücke zwischen *knows how* und *does* als Performanz (*shows how/performance*). Performanzorientierte Prüfungen in diesem Sinne bilden typische berufliche Handlungen in komplexitätsreduzierten Settings ab, z. B. in Form von rollenspielbasierten Simulationen (*approximations of practice*, Grossman et al., 2009). Sie bieten damit die Möglichkeit, Kernpraktiken des Lehrer*innenberufs, sogenannte *Core Practices* (s. Grossman et al., 2009), isoliert und standardisiert zu prüfen. In diesem Sinne können zum Beispiel das Führen eines Feedbackgesprächs, das Erklären eines physikalischen Phänomens oder auch das Lösen eines Konflikts zwischen Schülerinnen als *Core Practices* verstanden werden, die durch performanzorientierte Prüfungsformate geprüft werden können.

In anderen professionsorientierten Studiengängen wie der Medizin sind solche Prüfungsformate seit Langem etabliert (Pierre & Breuer, 2018). Hier kommen bspw. sogenannte *Objective Structured Clinical Examinations* (OSCE) zum Einsatz, die durch zumeist 10-12 etwa 5-minütige Prüfungsstationen, die als zusammenhängender Parcours absolviert werden, auf die Prüfung einer breiten Palette klinischer Kompetenzen abzielen (Sopka et al., 2018). An jeder Station steht dabei die Bewältigung einer spezifischen klinischen Aufgabe im Mittelpunkt, oft unterstützt durch z. B. Simulationspatient*innen, Computersimulationen oder medizinische Dokumente. Analog zu OSCE werden in der medizinischen Ausbildung außerdem *Objective Structured Teaching Examinations* (OSTE) eingesetzt, die die Prüfung von Lehr- und Kommunikationsfähigkeiten von Ausbilder*innen zum Ziel haben (Fakhouri & Nunes, 2019). Im Lehramtsstudium sind vergleichbare

performanzorientierte Prüfungsformate bisher kaum etabliert. Zu Übungszwecken kommen jedoch bereits simulationsbasierte Ansätze mit Fokus auf verschiedene *Core Practices* zum Einsatz (z. B. Führen von Elterngesprächen, Dotger et al., 2010; Gerich & Schmitz, 2016; Diagnosefähigkeiten, Fischer & Opitz, 2022). Vereinzelt existieren auch simulationsbasierte Assessmentformate, die jedoch bisher vor allem zu Forschungszwecken eingesetzt werden (z.B. für Erklärfähigkeiten, Kulgemeyer & Riese, 2018; Diagnosefähigkeiten, Kaiser & Möller, 2017).

Dieser Beitrag stellt daher die Entwicklung und Erprobung eines simulationsbasierten, performanzorientierten Prüfungsformates angelehnt an OSTE aus der Medizinausbildung für das Lehramtsstudium vor. Anhand von Ergebnissen aus ersten begleitenden Studien zur Durchführung des entwickelten OSTE-Prototypen soll zudem die Frage beantwortet werden, wie Studierende den Prototypen hinsichtlich der Eignung als Prüfungsformat für das Lehramtsstudium beurteilen.

Ein OSTE-Prototyp für das Lehramtsstudium im Fach Physik

Der OSTE-Prototyp wurde mit dem Ziel entwickelt, performative Aspekte professioneller Handlungskompetenz (angehender) Physiklehrkräfte (Blömeke, Gustafson & Shavelson, 2015) zu erfassen. Angelehnt an die vier Kompetenzbereiche Unterrichten, Beurteilen, Erziehen und Innovieren (KMK, 2024) bilden die insgesamt sieben 10- bis 20-minütigen Prüfungsstationen unterschiedliche, typische Anforderungen des Lehrkräfteberufs ab, wobei neben Neuentwicklungen bestehende Vorarbeiten adaptiert wurden (z. B. Feser, 2019; Kulgemeyer & Riese, 2018). Im Sinne des *Core-Practices*-Ansatz wird an jeder Prüfungsstation eine einzelne isolierte Kompetenz(-facette) erfasst. Dazu zählen: Unterrichtsplanung, Unterrichtsreflexion, Erklären eines physikalischen Phänomens (alle Kompetenzbereich Unterrichten), Beratung eines Schülers, Beurteilung schriftlicher Aufgabenlösungen (beide Kompetenzbereich Beurteilen), Moderation eines Konfliktgesprächs (Kompetenzbereich Erziehen) und kollegiale Unterrichtsentwicklung (Kompetenzbereich Innovieren). Der Prototyp umfasst sowohl eher bildungswissenschaftlich ausgerichtete Stationen als auch eher fachdidaktisch ausgerichtete Stationen. Zudem unterscheidet sich der Bearbeitungsmodus der Stationen: So enthält der OSTE neben Gesprächssimulationen mit Schauspieler*innen auch simulierte Videokonferenzen am Laptop sowie komplett schriftlich zu bearbeitende Stationen. Alle Stationen werden nacheinander von den Prüflingen durchlaufen, sodass ein kompletter Prüfungsparcours inklusive kurzer Pausen etwa 90 Minuten umfasst.

Wahrnehmung des OSTE-Prototypen durch Studierende

Mit dem Ziel der Erprobung und Weiterentwicklung des OSTE-Prototypen wurde der Prüfungsparcours im Sommersemester 2024 ganz oder teilweise mit insgesamt sechs Studierendengruppen an drei Universitätsstandorten durchgeführt. Je Erprobung absolvierten jeweils vier bis 11 Lehramtsstudierenden des Unterrichtsfachs Physik ($n_{\text{Physik}} = 34$) den gesamten Prüfungsparcours; jeweils 7 bis 12 Studierende in bildungs- bzw. erziehungswissenschaftlichen Seminaren ($n_{\text{BiWi}} = 19$) testeten ausschließlich die neuentwickelte Station zur Moderation eines Konfliktgesprächs.

Methodisches Vorgehen

Die Teilnehmenden in beiden Stichproben bearbeiteten jeweils nach Beenden einer Station einen Kurzfragebogen mit fünf-stufigen Likert-Items zu ihrer Wahrnehmung der

entsprechenden Station hinsichtlich Authentizität („Ich finde die dargestellte Situation authentisch.“, $0.708 < \alpha < 0.852$) und kognitiver Involviertheit („Ich habe mich in die Situation hineinversetzt.“, $0.724 < \alpha < 0.851$). In der Gesamtdurchführungsstichprobe füllten die Teilnehmenden zusätzlich einen Fragebogen zu demographischen Angaben, Erfahrungen mit für das Absolvieren des OSTE relevanten Unterrichtsinhalten sowie einer Gesamteinschätzung des OSTE-Prototyps am Ende des Prüfungsparcours aus. Hierbei kamen sechs-stufige Likert-Items zum Einsatz. Einige Studierende erklärten sich zudem zu einem vertiefenden Leitfadeninterview bereit. Die Daten aus den Fragebögen wurden deskriptiv analysiert. Die Leitfadeninterviews wurden mithilfe der qualitativen Inhaltsanalyse (Mayring, 2022) ausgewertet.

Ergebnisse und Ausblick

Die Ergebnisse deuten insgesamt auf eine positive Einschätzung des OSTE sowie der Konfliktstation in Bezug auf die Eignung als mögliches Prüfungsformat hin, zugleich aber auch auf Herausforderungen aufgrund des noch unbekanntes Formats. Im Detail waren die Teilnehmenden der Ansicht, dass der Prüfungsparcours authentische Anforderungen des Lehrer*innenberufs abbildet und prüft, was für den späteren Beruf als Lehrkraft relevant ist. Auch die Wichtigkeit der geprüften Inhalte für die gesamte Lehramtsausbildung wird als hoch eingeschätzt. Weniger positiv beurteilen die Teilnehmenden hingegen die Möglichkeit, sich gut auf einen solchen Prüfungsparcours vorbereiten und die Vorbereitung gut planen zu können sowie die Vergleichbarkeit der Leistungen von Studierenden im Prüfungsparcours. Die qualitativen Ergebnisse aus den leitfadengestützten Interviews unterstreichen diesen Eindruck insgesamt.

Tab. 5: Gesamteinschätzung des OSTE-Prototyps

Items	M	SD
In so einem Prüfungsparcours wird geprüft, was für meinen späteren Beruf als Lehrkraft relevant ist.	5.35	0.73
Auf so einen Prüfungsparcours kann man sich gut vorbereiten.	3.62	1.07
In so einem Prüfungsparcours sind die Leistungen von Studierenden gut vergleichbar.	3.76	0.99
So ein Prüfungsparcours bildet authentische Anforderungen des Lehrer*innenberufs ab.	5.44	0.71
Die Vorbereitung auf so einen Prüfungsparcours kann man gut planen.	3.82	1.03
In so einem Prüfungsparcours wird geprüft, was für die gesamte Lehramtsausbildung wichtig ist.	5.29	1.03

Insgesamt scheint der OSTE von den Studierenden als Prüfungsformat akzeptiert zu werden. Der Prototyp wird dabei insbesondere als authentisch in Bezug auf die beruflichen Anforderungen im Lehrkräfteberuf eingeschätzt. Offen ist unter anderem noch die Frage, aus welchen Gründen die Teilnehmenden die Vorbereitung auf einen solchen Prüfungsparcours als problematisch einschätzen (bspw., weil passende Lerngelegenheiten in vorherigen Kursen fehlen). Der Prototyp soll daher mit weiteren Studierenden an unterschiedlichen Universitätsstandorten erprobt werden, um eine belastbarere Datengrundlage und weitere Einschätzungen für die Weiterentwicklung zu gewinnen. Perspektivisch ist auch eine Übertragung auf andere Studienfächer möglich.

Literatur

- Biggs, J. & Tang, C. (2011). *Teaching for Quality Learning at University. What the students does* (4th Edition). Open University Press.
- Blömeke, S., Gustafsson, J.-E. & Shavelson, R. J. (2015). Beyond Dichotomies: Competence viewed as continuum. *Zeitschrift für Psychologie*, 223(1), 3–13. <https://doi.org/10.1027/2151-2604/a000194>
- Braun, E. (2021). Performance-based assessment of students' communication skills. *International Journal of Chinese Education*, 10(1), 221258682110062. <https://doi.org/10.1177/22125868211006202>
- Carless, D., Bridges, S. M., Chan, C. K. Y., & Glofcheski, R. (Eds.). (2017). *Scaling up assessment for learning in higher education*. Springer.
- Dotger, B. H., Dotger, S. C. & Maher, M. J. (2010). From Medicine to Teaching: The Evolution of the Simulated Interaction Model. *Innovative Higher Education*, 35(3), 129–141. <https://doi.org/10.1007/s10755-009-9128-x>
- Fakhouri, S. A., & Nunes, M. D. P. T. (2019). Objective structured teaching examination (OSTE): an underused tool developed to assess clinical teaching skills. A narrative review of the literature. *Sao Paulo Medical Journal*, 137, 193-200.
- Feser, M. S. (2019). *Physiklehrkräfte korrigieren Schülertexte*. Logos Verlag Berlin.
- Fischer, F. & Opitz, A. (Hrsg.). (2022). *Learning to Diagnose with Simulations*. Cham: Springer International Publishing. <https://doi.org/10.1007/978-3-030-89147-3>
- Flores, M. A., Veiga Simão, A. M., Barros, A. & Pereira, D. (2015). Perceptions of effectiveness, fairness and feedback of assessment methods: a study in higher education. *Studies in Higher Education*, 40 (9), 1523-1534.
- Gerich, M. & Schmitz, B. (2016). Using Simulated Parent-Teacher Talks to Assess and Improve Prospective Teachers' Counseling Competence. *Journal of Education and Learning*, 5(2), 285–301. <https://doi.org/10.5539/jel.v5n2p285>
- Grossman, P., Compton, C., Igra, D., Ronfeldt, M., Shahan, E., & Williamson, P. W. (2009). Teaching practice: A cross-professional perspective. *Teachers college record*, 111(9), 2055-2100.
- KMK (2024). Standards für die Lehrerbildung: Bildungswissenschaften. Beschluss der Kultusministerkonferenz vom 16.12.2004 i. d. F. vom 07.10.2022. URL: https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2004/2004_12_16-Standards-Lehrerbildung.pdf
- Kulgemeyer, C., & Riese, J. (2018). From professional knowledge to professional performance: The impact of CK and PCK on teaching quality in explaining situations. *Journal of Research in Science Teaching*, 55(10), 1393-1418.
- Miller, G. E. (1990). The assessment of clinical skills/competence/performance. *Academic medicine journal of the Association of American Medical Colleges*, 65(9), S63-7. <https://doi.org/10.1097/00001888-199009000-00045>
- Nicol, D. J. & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: a model and seven principles of good feedback practice. *Studies in Higher Education*, 31 (2), 199-218.
- St. Pierre, M. S., & Breuer, G. (Hrsg.) (2018). *Simulation in der Medizin: Grundlegende Konzepte - Klinische Anwendung* (2. Aufl.). Springer.
- Sopka, S., Simon, M., & Beckers, S. K. (2018). „Assessment drives learning “: Konzepte zur Erfolgs-und Qualitätskontrolle. In St. Pierre, & G. Breuer (Hrsg.), *Simulation in der Medizin: Grundlegende Konzepte-Klinische Anwendung* (2.Aufl.) (S. 97-109). Springer.
- Zeichner, K. (2010). Rethinking the connections between campus courses and field experiences in college-and university-based teacher education. *Journal of teacher education*, 61(1-2), 89-99.