

Entwicklung eines PCK-Tests zum experimentgestützten Chemieunterricht

Ziel

Ziel des Projekts ist es, die Wirksamkeit einer Lehrerfortbildung zu evaluieren, die auf die Optimierung der Qualitätsmerkmale von experimentgestütztem Chemieunterricht fokussiert.

Theoretischer Hintergrund

Die erziehungswissenschaftliche Forschung zur Unterrichtsqualität weist im Wesentlichen allgemeine bzw. fachunspezifische Unterrichtsmerkmale aus, die positiv mit den gemessenen Leistungs-, Interessens-, und Einstellungsvariablen der Schülerinnen und Schüler korrelieren (Clausen, 2002; Helmke, 2006). Aus der fachdidaktischen Forschung liegen für den Chemie- und Physikunterricht Erkenntnisse vor, dass hier insbesondere Merkmale der Zielorientierung, wie das Explizieren von Lernzielen und die Sequenzierung von Unterricht in überschaubare und aufeinander aufbauende Einheiten wichtig sind (Seidel et al., 2006). Aber auch Merkmale des Erkenntnisgewinnungsprozesses mit Hilfe von Experimenten, wie Entwicklung von Fragestellungen und selbstständige Versuchsplanung, wirken auf den Lernzuwachs bei Schülerinnen und Schülern (Schulz, 2011). Tesch und Duit (2004) konnten zeigen, dass die Leistungsentwicklung der Schülerinnen und Schüler positiv mit der Gesamtdauer der Experimentalphase korreliert, nicht aber mit der Dauer der Versuchsdurchführung. Das legt den Schluss nahe, dass für den Lernzuwachs eine strukturierte Einbettung mit entsprechender Vor- und Nachbereitung des Experiments notwendig ist. Aus den zuvor genannten Videostudien (Schulz, 2011; Seidel et al., 2006; Tesch & Duit, 2004), die im regulären Unterricht an deutschen Schulen angefertigt wurden, lässt sich ableiten, dass die auf ein Lernziel bezogene Strukturierung nach sachlogisch-inhaltlichen und funktional-lernprozessorientierten Gesichtspunkten, die auch als Sequenzierung bezeichnet werden kann, eine zentrale Rolle spielt. Insgesamt waren aber die Ausprägungen der entsprechenden Merkmale auf mittlerem Niveau (Schulz, 2011) bzw. noch darunter, sodass für den Chemie- und Physikunterricht in Deutschland auch im Vergleich zu anderen europäischen Ländern, wie z. B. Finnland oder die Schweiz, diesbezüglich ein Entwicklungsbedarf zu bestehen scheint (Börlin & Labudde, 2014).

Die oben angesprochene Entwicklung des Unterrichts mit Hinblick auf die Qualitätsmerkmale findet aber nicht von allein statt, sondern es müssen im Rahmen der zweiten und auch weiterhin in der dritten Phase der Lehrerbildung Lerngelegenheiten für diese Thematik geschaffen werden. Um an dieser Stelle eine fortlaufende Professionalisierung von Lehrkräften in der dritten Phase zu erreichen, muss nach Radtke (1996) das Handeln u. a. mit fachdidaktischem Wissen vernetzt und weiterentwickelt werden. Schmitt (2016) konnte im Rahmen einer eintägigen Fortbildung zur Förderung des experimentgestützten Erkenntnisgewinnungsprozesses zeigen, dass diese zu kurz- und mittelfristigen Veränderungen in der Einstellung der Lehrerinnen und Lehrern gegenüber dem Fortbildungsinhalt wie auch im fachdidaktischen Wissen zu Experimenten führt. Auch das Coaching mit Videomaterial aus dem eigenen Unterricht der Lehrkräfte hat sich in mehreren Studien bewährt (Schulz, 2011; Seidel et al., 2006; Trendel, Wackermann & Fischer, 2007). Schulz (2011) konnte unter quasi-experimentellen Bedingungen in einem stundenspezifischen Einzelcoaching, in dem Videos aus dem Unterricht der jeweiligen Lehrkräfte Anwendung fanden, zeigen, dass dieses bei erneuter Durchführung zu einer verstärkten Merkmalausprägung und in Folge auch zu einem höheren Lernzuwachs bei den Schülerinnen und Schülern führt. Die Lernwirksamkeit eines unterrichtsthemenspezifischen Coachings zu

Qualitätsmerkmalen experimenteller Phasen auf der Handlungsebene der Lehrerinnen und Lehrer und auf der kognitiven und affektiven Ebene der Schülerinnen und Schüler ist demnach belegt.

Ungeklärt ist bisher, ob diese Effekte in Bezug auf die Qualitätsmerkmalsausprägung auch im Rahmen einer nicht unterrichtsthemenspezifischen Gruppenfortbildung erreicht werden können.

Forschungsfragen

Aus dem Ziel und dem theoretischen Hintergrund des Projekts lässt sich folgende Forschungsfrage für die Hauptstudie ableiten:

Führt eine Lehrerfortbildung, die auf die Qualitätsmerkmale von Chemieunterricht mit experimentellen Phasen fokussiert, bei den teilnehmenden Lehrkräften zu Veränderungen der epistemischen Überzeugungen, des fachdidaktischen Wissens und des unterrichtspraktischen Handelns?

Um diese beantworten zu können, müssen zunächst geeignete Testinstrumente entwickelt und pilotiert werden. Daraus ergibt sich für die Pilotierung folgende Forschungsfrage:

Sind die entwickelten Instrumente objektiv, reliabel und valide zur Messung der epistemischen Überzeugungen und des fachdidaktischen Wissens der Lehrkräfte?

Methoden und Design der Pilotstudie

Zur Erfassung der Lehrerkognition wurden zwei Testinstrumente eingesetzt, ein Instrument zu epistemischen Überzeugungen und eines zum fachdidaktischen Wissen der Lehrkräfte zu Experimenten und deren methodischer Umsetzung.

Für das Instrument zur Erfassung des fachdidaktischen Wissens wurden 26 Items mit je vier Aussagen in Anlehnung an Tepner & Dollny (2014) konstruiert. Jede dieser Aussagen musste auf einer vierstufigen Likert-Skala bewertet werden. Zusätzlich erfolgte eine Validierung der Items zum fachdidaktischen Wissen durch ein Expertenrating, in dem Hochschullehrende der Chemiedidaktik um ihre Einschätzung gebeten wurden.

Für das Instrument zu den epistemischen Überzeugungen wurden 17 Items mit jeweils einer Aussage aus dem Instrument von Lamprecht (2011) ausgewählt und adaptiert. Dabei wurden die Items ausgewählt, die er für eine Zuordnung der Lehrkräfte zu einem der drei Überzeugungstypen (Training, Diskursiv, Vermittlung) genutzt hat (Lamprecht, 2011). Jede Aussage musste auf einer fünfstufigen Likert-Skala bewertet werden.

Die Lehrkräfte der Pilotierungsstichprobe ($N = 23$) wurden online mittels LimeSurvey mit einer festen Reihenfolge der Items getestet. Dabei wurden nach einem Demographieteil zunächst Items zum fachdidaktischen Wissen und dann die Items zu den Überzeugungen abgefragt. Die Probanden mussten jede Aussage bewerten (forced-choice) um fortfahren zu können.

Ergebnisse der Pilotstudie

Im Test zum fachdidaktischen Wissen ergibt sich für das Expertenrating ($N = 3$) bei Nutzung einer vierstufigen Likert-Skala nur eine mittelmäßige Übereinstimmung ($\kappa_{\text{Fleiss}} = .45$), gemessen auf Aussagenebene über alle 104 Aussagen. Bei einer dichotomen Kodierung, bei der die beiden (eher) zustimmenden und die beiden (eher) ablehnenden Abstufungen zusammengefasst wurden, verbessert sich die Interraterreliabilität ($\kappa_{\text{Fleiss}} = .67$) unter Einbeziehung aller Aussagen. Durch Eliminierung der Aussagen, bei denen das Expertenrating sehr breit streute, konnte die Interraterreliabilität bei den übrig gebliebenen 72 Aussagen für die vierstufige Kodierung verbessert ($\kappa_{\text{Fleiss}} = .62$) und für die dichotome Kodierung eine gute Interraterreliabilität erreicht werden ($\kappa_{\text{Fleiss}} = .91$), ohne die Validität des Messinstruments substanziell zu senken.

Die Ergebnisse der befragten Lehrkräfte ($N = 23$) wurden einer Rasch-Skalierung unterzogen. Unabhängig von der einbezogenen Anzahl der Aussagen und der Art der Kodierung ergibt sich aber eine unzureichende EAP-Reliabilität ($Rel_{EAP} < .60$).

Für das Testinstrument ergibt sich in der Datenanalyse ein zweidimensionales Modell (vgl. Tabelle 1). Die erste Dimension bilden zehn Items zum Chemieunterricht, die zweite sieben Items zum Lernen von Chemie in der Schule.

Tabelle 1: Vergleich des ein- und zweidimensionalen Modells

Modell	Deviance	AIC	chisq	df	p
eindimensional	936.62	1199.99	33.96	2	0
zweidimensional	902.66	1172.31			

Für die beiden Dimensionen ergeben sich mittlere bis gute EAP-Reliabilitäten und akzeptable Infit-Werte. Dabei ist auffällig, dass die Varianz zwischen den beiden Dimensionen stark unterschiedlich groß ist (vgl. Tabelle 2).

Tabelle 2: EAP-Reliabilitäten, Varianzen und Infit-Werte für die beiden Dimensionen

Dim1: Chemieunterricht			Dim2: Lernen von Chemie in der Schule		
Rel_{EAP}	Varianz	Infit	Rel_{EAP}	Varianz	Infit
.862	2.325	0.58 – 1.24	.602	0.162	0.86 – 1.50

Konsequenzen für die Hauptstudie

Bis zum Einsatz in der Hauptstudie ist eine Erhöhung der Stichprobengröße sowohl für die Experten wie auch bei den Lehrkräften erforderlich, da einzelne (schlechte) Kennwerte der Test möglicherweise auch mit der geringen Stichprobengröße zusammenhängen. Zudem wird geprüft, ob es sich bei dem Konstrukt, das dem Testinstrument zum fachdidaktischen Wissen zugrunde liegt, gegebenenfalls um ein mehrdimensionales handelt. In jedem Fall ist eine Überarbeitung des Instruments zum fachdidaktischen Wissen zwingend erforderlich und für das Instrument zu den epistemischen Überzeugungen mindestens angeraten. Möglicherweise bieten die Performanztests zur Planung und Reflexion von Unterricht aus dem Projekt ProfileP+ (Kulgemeyer et al., 2019) hier eine Alternative zum hier dargestellten Test zur Überprüfung des fachdidaktischen Wissens der Lehrkräfte.

Ausblick auf die Hauptstudie

Die eintägige Fortbildungsmaßnahme mit insgesamt sechs Stunden Interventionszeit besteht aus einem instruktionalen und einem praktischen Teil, bei dem Videos aus dem eigenen Unterricht der teilnehmenden Lehrkräfte zum Einsatz kommen. Zur Evaluation der Wirksamkeit der Lehrerfortbildung werden Daten zu den vier Ebenen der Wirksamkeit nach Lipowsky (2010) erhoben.

Die erste ist die affektive Ebene der Lehrkräfte. Sie wird im Rahmen des Projekts aus den Reaktionen und Einschätzungen der Lehrkräfte zur Fortbildung und deren Inhalt gebildet und durch einen Fragebogen nach Schmitt (2016) erhoben.

Die zweite Ebene betrifft die Erweiterung der Lehrerkognition. Sie wird in diesem Rahmen aus den epistemischen Überzeugungen und als fachdidaktisches Wissen zu Experimenten und deren methodischer Umsetzung als Teile professioneller Kompetenz (Baumert & Kunter, 2011) gebildet.

Das unterrichtspraktische Handeln der Lehrkräfte auf der dritten Ebene wird auf Grundlage der Videoanalyse zweier Unterrichtsstunden (prä-post) bewertet. Dabei kommt das Kategoriensystem zum Einsatz, das von Schulz (2011) entwickelt und verwendet wurde.

Auf der vierten Ebene werden die Effekte auf Schülerinnen und Schüler als Wahrnehmung der Unterrichtsqualität mittels Fragebogen (Schulz, 2011) erhoben.

Literatur

- Baumert, J., & Kunter, M. (2011). Das Kompetenzmodell von COACTIV. In M. Kunter, J. Baumert, W. Blum, & M. Neubrand (Hrsg.), *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV* (S. 29–54). Münster: Waxmann.
- Börlin, J., & Labudde, P. (2014). Practical Work in Physics Instruction: An Opportunity to learn? In H. E. Fischer, P. Labudde, K. Neumann, & J. Viiri (Hrsg.), *Quality of Instruction in Physics. Comparing Finland, Switzerland and Germany* (S. 111–128). Münster: Waxmann.
- Clausen, M. (2002). *Unterrichtsqualität: Eine Frage der Perspektive (Pädagogische Psychologie und Entwicklungspsychologie: Bd. 29)*. Münster: Waxmann.
- Helmke, A. (2006). Was wissen wir über guten Unterricht? Über die Notwendigkeit einer Rückbesinnung auf den Unterricht als dem "Kerngeschäft" der Schule (II.Folge). *Zeitschrift für Pädagogik*, 58 (2), 42-45.
- Kulgemeyer, C., Vogelsang, C., Borowski, A., Buschhüter, D., Enkrott, P., Kempin, M., Reinhold, P., Riese, J., Schecker, H., & Schröder, J. (2019). Entwicklung von Professionswissen und Unterrichtsperformanz im Lehramtsstudium Physik. *Zeitschrift für Pädagogik*, 71 (4), 473-492.
- Lamprecht, J. (2011). *Ausbildungswege und Komponenten professioneller Handlungskompetenz: Vergleich von Quereinsteigern mit Lehramtsabsolventen für Gymnasien im Fach Physik*. Berlin: Logos.
- Lipowsky, F. (2010). Lernen im Beruf – Empirische Befunde zur Wirksamkeit von Lehrerfortbildung. In F. Müller, A. Eichenberger, M. Lüders & J. Mayr (Hrsg.), *Lehrerinnen und Lehrer lernen – Konzepte und Befunde zur Lehrerfortbildung* (S. 51-72). Münster: Waxmann.
- Oser, F. K., & Baeriswyl, F. J. (2001). Choreographies of Teaching: Bridging Instruction to Learning. In V. Richardson (Hrsg.), *Handbook of research on teaching* (S. 1031–1065). Washington, DC: AERA.
- Radtke, F.-O. (1996). *Wissen und Können – die Rolle der Erziehungswissenschaft in der Erziehung*. Opladen: Leske + Buderich.
- Schmitt, A. K. (2016). *Entwicklung und Evaluation einer Chemielehrerfortbildung zum Kompetenzbereich Erkenntnisgewinnung*. Berlin: Logos.
- Schulz, A. (2011). *Experimentierspezifische Qualitätsmerkmale im Chemieunterricht: Eine Videostudie*. Berlin: Logos.
- Seidel, T., Prenzel, M., Rimmel, R., Schmidt, K., Kobarg, M., & Herweg, C., Dalehefte, I. M. (2006). Unterrichtsmuster und ihre Wirkungen: Eine Videostudie im Physikunterricht. In M. Prenzel & L. Allolio-Näcke (Hrsg.), *Untersuchungen zur Bildungsqualität von Schule. Abschlussbericht des DFG-Schwerpunktprogramms BiQua*. Münster: Waxmann.
- Tesch, M. & Duit, R. (2004). Experimentieren im Physikunterricht – Ergebnisse einer Videostudie. *Zeitschrift für Didaktik der Naturwissenschaften*, 18, 7-28.
- Tepner, O., & Dollny, S. (2014). Entwicklung eines Testverfahrens zur Analyse fachdidaktischen Wissens. In D. Krüger, I. Parchmann, & H. Schecker (Hrsg.), *Methoden in der naturwissenschaftsdidaktischen Forschung* (S. 311–323). Berlin: Springer.