

Fabian Poensgen¹
 Christiane S. Reiners¹

¹Universität zu Köln

Modellierung und Diagnose experimenteller Kompetenzen in der Chemielehrer*innenbildung

Experimentelles Arbeiten im Labor ist für Studierende ein wesentliches Charakteristikum ihres Chemiestudiums und nimmt auch unter Lehrenden einen enormen Stellenwert ein: „In einem sind sich die Chemiefachbereiche einig: Chemie studieren im Fernunterricht ohne Laborpraktika ist sinnlos“ (Zbikowski, 2020). In der Forschung ist der Erfolg ressourcenintensiver Laborpraktika jedoch umstritten und deren Ziel – die Entwicklung experimenteller Kompetenzen – wenig untersucht (Abd-El-Khalick et al., 2004; Bretz, 2019; Heidrich, 2017; Pullen, Thickett & Bissember, 2018). In dem hier vorgestellten Forschungsprojekt soll dies anhand von Chemielehramtsstudierenden der Universität zu Köln im ersten bis vierten Semester geschehen. Dazu bietet es sich nach Bretz (2012) an, zwei grundlegende Fragen zu ergründen: „Was sollen die Studierenden wissen?“ und „wie können wir heraus finden, dass sie es wissen?“ Diese Fragen bedingen sich jedoch gegenseitig, da für rein normativ vorgegebene Kompetenzziele nicht unbedingt auch adäquate Prüfverfahren (vgl. Bernholt, 2010) entwickelt werden können und damit gleichsam überlegt werden muss: „Was kann als Beleg dienen, dass Studierende etwas über Chemie wissen?“ (Bretz, 2012). Experimentelle Kompetenzen werden auf diese Weise erst über spezifische Aufgaben- oder Problemstellungen erfahrbar (Gut-Glanzman, 2012) und sind damit kontextabhängig (Blömeke, Gustafsson & Shavelson, 2015).

Aus diesem Grund wurde hier im Sinne einer formativen Evaluation ein iterativ helikales Vorgehen mit sich zuspitzenden Rückkopplungs-Schleifen gewählt, in dem ein auf Literaturlösungsbasis kompiliertes Kompetenzmodell und ein dazu konzipiertes Prüfverfahren zur Diagnose experimenteller Kompetenzen kontinuierlich empirisch überprüft und weiterentwickelt wurden.

Modellierung experimenteller Kompetenzen

Experimentelle Kompetenzen werden als ein von anderen Kompetenzbereichen unabhängig messbares (Kompetenz-)Konstrukt betrachtet (Gut & Mayer, 2018), das sowohl mögliches Ergebnis als auch Voraussetzung für sinnvolle Laborarbeit ist (Abd-El-Khalick et al., 2004). Ziel der Kompetenzentwicklung soll es nach Pullen et. al. (2018) sein, dass die Studierenden ein fundamentales Verständnis chemischer Prinzipien und Theorien entwickeln und damit Sachverhalte erklären („K“), chemische Konzepte zur Problemlösung anwenden („P“), fundamentale Experimentiertechniken nutzen („F“) sowie verantwortungsvoll und sicher experimentieren („S“). Experimentelle Kompetenzen werden hier deshalb in Anlehnung an die Formulierung Weinerts (2002) verstanden als die bei Chemielehramtsstudierenden verfügbaren oder durch sie erlernbaren prozessbezogenen („P“) und konzeptbezogenen („K“) kognitiven Fähigkeiten und psychomotorischen Fertigkeiten („F“), um die experimentelle Methode in variablen Situationen sicher, verantwortungsvoll („S“) und reflektiert anwenden zu können, d.h. die Formulierung von Problem- bzw. Fragestellungen sowie Hypothesen und das Planen, Durchführen und Auswerten von Experimenten, sowie die damit verbundenen motivationalen, volitionalen und sozialen Bereitschaften und Fähigkeiten.

Auf Grundlage einer umfangreichen Literaturrecherche wurden etwa 100 Kompetenzaspekte experimenteller Kompetenz identifiziert, die auf vier Niveaustufen den in der Definition heraus gearbeiteten Kompetenzdimensionen zugeordnet wurden. Anhand von videographierten Experimentiersituationen, den Erkenntnissen aus dem unten dargestellten Prüfverfahren und einer Expertenbefragung wurde das Modell hinsichtlich Stufung und Kategorisierung mehrfach evaluiert und grundlegend überarbeitet (vgl. Bernholt, 2010).

Übergeordnetes Ziel der Modellierung ist es, einen validen, empirisch abgesicherten Orientierungsrahmen zur Kompetenzentwicklung im Laborpraktikum zu entwickeln, um für Lernende und Lehrende Transparenz herzustellen und eine differenzierte Diagnose als Grundlage für individuelle Förderung zu ermöglichen.

Diagnose experimenteller Kompetenzen

Kompetenzen sind latente Merkmale von Menschen und können somit nicht direkt beobachtet, sondern nur aus erkennbaren Handlungen geschlussfolgert werden (Blömeke et al., 2015, S. 2). Obwohl experimentelle Kompetenzen zumeist in Form von schriftlichen Test, der Analyse von Laborprotokollen oder Computersimulationen erhoben werden, herrscht weitgehende Einigkeit, dass videographierte Hands-On Tests als Referenz anzusehen sind (Kambach, 2018). Gerade im Hochschulbereich liegt ein Fokus auf den manuellen Fertigkeiten, die nur durch die tatsächliche Durchführung erfahrbar sind (Meier, 2016), wofür jedoch kein geeignetes Prüfungsformat vorliegt (Heidrich, 2017).

Erhebung

Deshalb wurde hier ein Prüfverfahren entwickelt, bei dem die Studierenden zu einem vorgegebenen, lebensweltlichen Kontext (Abd-El-Khalick et al., 2004) wie z.B. Glühwein oder Essiggurkenwasser eine Fragestellung entwickeln, Hypothesen formulieren, ein Experiment entwerfen, im Labor durchführen und schriftlich auswerten. Insofern ermöglicht das weitgehend geöffnete Prüfungsdesign (Baur, Hummel, Emden & Schröter, 2020) genuin die Formulierung von Fragestellungen als konstitutivem Bestandteil experimenteller Kompetenz (Fricke, 2018), Indikator für Fachwissen (Ströker nach Reiners, 2017) und Ausgangs- sowie Kumulationspunkt für die Suche im Hypothesen- und Experimentier-Raum (Klahr & Dunbar, 1988). Zudem stellt das formative Prüfverfahren diagnostisch den Problemlöseprozess in den Mittelpunkt (Meier, 2016) und lässt stark divergierende Lösungsansätze zu, die anstelle der Aufgabenschwierigkeit das Kompetenzniveau determinieren.

Während der praktischen Arbeit tragen die Studierenden Kamerabrillen (Chalfen, 2014) und werden von Mitarbeiter*innen begleitet. Letztere tragen ebenfalls Kamerabrillen, bieten Hilfestellungen an, achten auf Sicherheitsaspekte und stellen den Studierenden an passenden Stellen vorgegebene Fragen im Rahmen eines Leitfrageninterviews. Auf diese Weise werden in einer qualitativen, nicht-experimentellen Längsschnittstudie mit explorativem Charakter (Döring & Bortz, 2016) anhand von drei Studierendekohorten ($3 \leq n \leq 15$) zu bisher acht Zeitpunkten Daten in Form von videographierten Handlungen, Äußerungen, Antworten auf die Fragen der Mitarbeitenden, schriftlichen Laborprotokolle und semesterbegleitenden Selbsteinschätzungen erhoben, so dass eine Methoden- und Daten-Triangulation (Döring & Bortz, 2016) möglich wird.

Auswertung

Die gewonnenen Daten werden mittels qualitativer Inhaltsanalyse nach Mayring (2015) auf Grundlage des Kompetenzmodells ausgewertet. Dieses wurde für die Selbsteinschätzung als Fragebogen operationalisiert (Cronbachs $\alpha = .940$, bei $n = 203$), während für die videographierten und schriftlichen Elemente Kodiermanuale erstellt wurden. Dazu wurden nach Fischer & Neumann (2012; basierend auf Jacobs, Kawanaka & Stigler) jeweils für die Handlungen, verbalen Äußerungen und Protokolle passende, gestufte Kategoriensysteme entwickelt. Diese wurden nach einer Probekodierung wiederum vom Kompetenzmodell ausgehend über das Prüfverfahren bis hin zum Kodiertraining im Hinblick auf die Gütekriterien Definitionen, Indikatoren, Ankerbeispiele, Training, und Kalibrierung (Hugener, Rakocy, Pauli & Reusser, 2006) in mehreren Schleifen in einem Wechselspiel aus induktivem und deduktivem Vorgehen (Brosius, Haas & Koschel, 2016) evaluiert und revidiert. Um die Zuverlässigkeit und Reliabilität der Ratings zu steigern, wurden die Ankerbeispiele im Hinblick auf drei unterschiedliche Problemtypen (Gut, Metzger, Hild & Tardent, 2014) differenziert. Bis auf eine Studierende führten die Studierenden bisher nämlich unabhängig von der Aufgabenstellung lediglich Titrationen, Destillationen oder Nachweisreaktionen durch.

Erste Ergebnisse

Obwohl noch Schritte zur internen und externen Validierung ausstehen, konnten Modell und Prüfverfahren so weit entwickelt werden, dass sich alle Kompetenzaspekte identifizieren lassen. Es zeigt sich jedoch, dass nicht alle Aspekte mit jedem Erhebungsinstrument gleichermaßen erfassbar sind, sondern erst die Kombination tiefergehende Einblicke in das Kompetenzniveau der Studierenden erlaubt. Die Verteilung der Kodierungen zu den vier Stufen lässt zudem den Schluss zu, dass das Modell trotz Anlehnung an Studien aus dem schulischen Bereich ausreichend komplex ist, was aber sicherlich auch durch das erheblich komplexitätssteigernde offene Prüfverfahren (Baur et al., 2020) unterstützt wird.

Bisher wurden sechs Prüfungen doppelt kodiert und ergaben in Anbetracht des hoch inferenten Kategoriensystems mit Cohens $\kappa \geq .59$ und prozentualer Übereinstimmung $\geq .70$ zufriedenstellende Übereinstimmungen (Fischer & Neumann, 2012). Auf Grund der variierenden Anzahl an Analyseeinheiten wurden die Reliabilitäten über ein gestuftes Verfahren ermittelt, bei dem die Identifikation kodierwürdiger Ereignisse nur als prozentuale Übereinstimmung berechnet werden kann.

Was den bislang diagnostizierten Stand der experimentellen Kompetenzen angeht, können Überschneidungen, aber auch Unterschiede zu vergleichbaren Studien konstatiert werden:

- Forschungsfragen und Hypothesen werden oft um spezifische experimentelle Methoden herum konstruiert und haben damit eher bestätigenden Charakter (Kambach, 2018).
- Der Umgang mit Variablen ist sowohl im Hinblick auf die Hypothesenbildung (vgl. Baur, 2018) als auch die Kontrolle (Meier, 2016; Schwichow & Nehring, 2018) oberflächlich.
- Im Bereich Sicherheit gelingt es den Studierenden häufig nicht, Gefahren und präventive Maßnahmen explizit zu benennen, selbst wenn sie letztere implizit korrekt ergreifen.
- Messungen werden teilweise ungenau durchgeführt, Beobachtungen und Daten nur entsprechend den Erwartungen ausgewertet (vgl. Chinn & Brewer, 1993).
- Studierende arbeiten jedoch anders als Schüler*innen (Baur, 2018) oft mit Blindproben.
- Im Gegensatz zu Schüler*innen (Meier, 2016) sind Studierende durchaus bereit, Hypothesen zu falsifizieren, auch wenn sie darin einen „Misserfolg“ sehen.

Literatur

- Abd-El-Khalick, F., BouJaoude, S., Duschl, R., Lederman, N., Mamlok-Naaman, R., Hofstein, A. et al. (2004). Inquiry in science education: International perspectives. *Science Education*, 88(3), 397–419.
- Baur, A. (2018). Fehler, Fehlkonzepte und spezifische Vorgehensweisen von Schülerinnen und Schülern beim Experimentieren. *Zeitschrift für Didaktik der Naturwissenschaften*, 24(1), 115–129.
- Baur, A., Hummel, E., Emden, M. & Schröter, E. (2020). Wie offen sollte offenes Experimentieren sein? Ein Plädoyer für das geöffnete Experimentieren. *MNU*, (2), 125–128.
- Bernholt, S. (2010). Kompetenzmodellierung in der Chemie. Theoretische und empirische Reflexion am Beispiel des Modells hierarchischer Komplexität 2010. Berlin: Logos-Verl.
- Blömeke, S., Gustafsson, J.-E. & Shavelson, R. J. (2015). Beyond Dichotomies: Competence viewed as a continuum. *Zeitschrift für Psychologie*, 223(1), 3–21.
- Bretz, S. L. (2012). Navigating the Landscape of Assessment. *JCE*, 89(6), 689–691.
- Bretz, S. L. (2019). Evidence for the Importance of Laboratory Courses. *JCE*, 96(2), 193–195.
- Brosius, H.-B., Haas, A. & Koschel, F. (2016). Methoden der empirischen Kommunikationsforschung. Wiesbaden: Springer VS.
- Chalfen, R. (2014). ‘Your panopticon or mine?’ Incorporating wearable technology’s Glass and GoPro into visual social science. *Visual Studies*, 29(3), 299–310.
- Chinn, C. A. & Brewer, W. F. (1993). The Role of Anomalous Data in Knowledge Acquisition. A Theoretical Framework and Implications for Science Instruction. *RER*, 63(1), 1–49.
- Döring, N. & Bortz, J. (2016). Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften. Springer Berlin Heidelberg.
- Fischer, H. E. & Neumann, K. (2012). Video Analysis as a Tool for Understanding Science Instruction. In D. Jorde & J. Dillon (Eds.), *Science Education Research and Practice in Europe*. Rotterdam: Sense, 114–139.
- Fricke, A. (2018). Interaktive Skripte Im Physikalischen Praktikum. Entwicklung und Evaluation Von Hypermedien Für Die Nebenfachausbildung. Berlin: Logos.
- Gut, C. & Mayer, J. (2018). Experimentelle Kompetenz. In D. Krüger, I. Parchmann & H. Schecker (Hrsg.), *Theorien in der naturwissenschaftsdidaktischen Forschung*. Berlin, Heidelberg: Springer 121–141.
- Gut-Glanzmann, C. (2012). Modellierung und Messung experimenteller Kompetenz. Berlin: Logos.
- Gut, C., Metzger, S., Hild, P. & Tardent, J. (2014). Problemtypenbasierte Modellierung und Messung experimenteller Kompetenzen. *PhyDiD B*.
- Heidrich, J. (2017). Erfassung von Experimentierkompetenz im universitären Kontext. Entwicklung und Validierung eines Experimentiertests zum Themenbereich Optik. Christian-Albrechts-Universität, Kiel.
- Hugener, I., Rakocy, K., Pauli, C. & Reusser, K. (2006). Videobasierte Unterrichtsforschung. In S. Rahm, I. Mammes & M. Schratz (Hrsg.), *Schulpädagogische Forschung Bd. 1*. Innsbruck: Studienverl., 41–53.
- Kambach, M. (2018). Experimentierprozesse von Lehramtsstudierenden der Biologie. Humboldt-Universität zu Berlin.
- Klahr, D. & Dunbar, K. (1988). Dual space search during scientific reasoning. *Cogn Sci*, 12(1), 1–48.
- Mayring, P. (2015). *Qualitative Inhaltsanalyse. Grundlagen und Techniken*. Weinheim: Beltz.
- Meier, M. (2016). Entwicklung und Prüfung eines Instruments zur Diagnose der Experimentierkompetenz von Schülerinnen und Schülern. Berlin: Logos.
- Pullen, R., Thickett, S. C. & Bissember, A. C. (2018). Investigating the viability of a competency-based, qualitative laboratory assessment model in first-year undergraduate chemistry. *CERP*, 19(2), 629–637.
- Reiners, C. S. (2017). *Chemie vermitteln*. Berlin, Heidelberg: Springer Berlin Heidelberg.
- Schwichow, M. & Nehring, A. (2018). Variablenkontrolle beim Experimentieren. *ZfDN*, 24(1), 217–233.
- Weinert, F. E. (2002). Vergleichende Leistungsmessungen in Schulen. In F. E. Weinert (Hrsg.), *Leistungsmessungen in Schulen*, S. 17–31. Weinheim: Beltz.
- Zbikowski, F. (2020). „Um die Erstsemester machen wir uns große Sorgen“. *Nachr. Chem.*, 68(10), 8–9.