

Netzwerke der Unterrichtsqualität - Sind Theorien praxisrelevant?

Theoretische Grundlagen

Unterrichtsqualitätsbeurteilungen sind ein zentrales Element der Lehrkräftebildung. Vor allem in der zweiten Phase nehmen sie eine wichtige Stellung ein und sollen angehenden Lehrkräften dabei helfen, ihren Unterricht zu optimieren. In der Praxis bestehen jedoch häufig Defizite in der Kommunikation, die sich vor allem zwischen, aber auch innerhalb der Gruppen von Fachleitungen und Referendar:innen aufzeigen lassen (Heinitz & Nehring, 2023). Für die zweite Phase der Lehrkräftebildung liegen nur wenige Vorgaben vor, sodass Unterschiede zwischen den Fachseminaren zu einem gewissen Maß absehbar sind (Weber & Czerwenka, 2021). Diese offenen Vorgaben bieten den Fachseminaren zwar die Möglichkeit, flexibel auf die Anforderungen der Referendar:innen zu reagieren, erschweren jedoch auch die Vergleichbarkeit zwischen den Seminaren und damit innerhalb der zweiten Phase der Lehrkräftebildung. Zusätzlich weisen auch Systematisierungen zur Erfassung der Unterrichtsqualität in der Forschung häufig Unterschiede in der Konzeptualisierung und Operationalisierung auf (Alp Christ et al., 2022). Zwar gibt es mit den drei Basisdimensionen (Klieme et al., 2001) eine (im deutschsprachigen Raum) weit verbreitete und genutzte Grundlage zur Erfassung der Unterrichtsqualität, allerdings äußert sich an unterschiedlichen Stellen ein Ergänzungsbedarf, der sowohl fachspezifische als auch generische Aspekte umfasst (Praetorius et al. 2020a). Zusätzlich kommt eine Diskussion um die theoretische Fundierung der Basisdimensionen hinzu (Rothland, 2024), die dazu beiträgt, dass es zwar eine potenzielle gemeinsame Grundlage für die Erfassung der Unterrichtsqualität gibt, diese jedoch an vielen Stellen unterschiedlich ausgelegt wird. Dieser Pluralismus, der sich bereits bei theoretischen Konstrukten zur Unterrichtsqualität zeigt, kann direkte Auswirkungen auf die Praxis haben, indem unterschiedliche Grundlagen genutzt oder auch dieselben Grundlagen unterschiedlich interpretiert werden und somit abweichende Voraussetzungen für angehende Lehrkräfte entstehen. Damit bleibt die Frage offen, ob theoretische Konstrukte zur Unterrichtsqualität überhaupt Eingang in die Praxis der Lehrkräftebildung finden.

Forschungsfragen

Bislang liegen nur wenige Erkenntnisse vor, inwiefern die Pluralität zugrundeliegender Theorien in der Praxis zu einer Hybridisierung von Systematisierungen der Unterrichtsqualität führt oder ob diese vergleichbar zur Literatur etabliert werden. Folgende Forschungsfragen wurden deshalb im Rahmen dieser Studie verfolgt:

1. Welche Unterrichtsqualitätsmerkmale werden bei einer offenen Beurteilung einer Unterrichtsstunde durch Fachleitungen und Referendar:innen genutzt?
2. Inwiefern weisen beide Gruppen Unterschiede in der Vernetzung von Unterrichtsqualitätsmerkmalen auf?
3. Inwiefern weisen die Vernetzungen der Unterrichtsqualitätsmerkmale Abweichungen zu theoretischen Grundlagen der Unterrichtsqualität auf?

Methoden

Grundlage für die Datenerhebung bildeten 34 Interviews (17 Fachleitungen und 17 Referendar:innen mit dem Fach Chemie). Im ersten Schritt wurde eine Videovignette aus dem

Chemieunterricht mit einer Länge von 25 min. angesehen, die dann in einem leitfadengestützten Interview mit Blick auf die Unterrichtsqualität beurteilt werden sollte. Den Teilnehmenden wurden für die Beurteilung keine Unterrichtsqualitätsmerkmale vorgegeben. Die Interviews hatten eine durchschnittliche Länge von 60 min. wobei die Gruppen keinen Unterschied in der durchschnittlichen Länge aufgewiesen haben. Anschließend wurden alle Interviews transkribiert und die Aussagen mit Hilfe des Kodiermanuals zum SEP-Framework (Heinitz & Nehring, 2023) kodiert (FF1). Dabei wurden insgesamt mehr als 2700 Aussagen erfasst, von denen wiederum 2201 Aussagen einen direkten Bezug zur Unterrichtsqualität aufwiesen und in die Analyse aufgenommen wurden. Zur Untersuchung der Vernetzung von Qualitätsmerkmalen wurde eine empirische Netzwerkanalyse durchgeführt und zunächst jedes Interview separat ausgewertet. Die Kodiereinheiten wurden dafür in ihrer natürlichen Gesprächsreihenfolge belassen und lediglich Füllsätze herausgenommen (kodiert als „Off-Topic“). Zwei Merkmale wurden als „vernetzt“ gewertet, wenn sie entweder in direkter Folge oder mit maximal einem Zwischeneinschub durch ein anderes Merkmal genannt wurden. Die einzelnen Vernetzungen wurden anschließend ausgezählt. Für die weitere Analyse wurden nur Vernetzungen berücksichtigt, die mehr als einmal auftraten. Im nächsten Schritt erfolgte eine Louvain-Clusteranalyse in R (Package: igraph), bei der zunächst alle Teilnehmenden gemeinsam betrachtet wurden, bevor getrennte Gruppenanalysen erfolgten (FF2). Da einige wenige Merkmale unverhältnismäßig häufig auftraten, wurden zusätzliche Analysen berechnet, bei denen die häufigsten Vernetzungen zwischen Clustern (Inter-Cluster-Vernetzungen) ausgeschlossen wurden, um den Einfluss auf die Modularität der jeweiligen Cluster zu überprüfen. In einem letzten Schritt wurden die Cluster den sieben Dimensionen des SEP-Frameworks gegenübergestellt, um Abweichungen zwischen Praxis und theoretischer Grundlage zu analysieren (FF3).

Ergebnisse

Bezüglich der Verwendung von Merkmalen (FF1) zeigt sich zunächst ein starker Fokus auf Dimension I des SEP-Frameworks („Auswahl und Thematisierung von Inhalten“). Am zweithäufigsten wird Dimension II („Kognitive Aktivierung“) verwendet, wobei hier ein Fokus auf die „Konstruktive Einbindung von eigenen Ideen und Schüler:innenvorstellungen“ gelegt wurde. Trotz einer vergleichbaren durchschnittlichen Interviewlänge, nennen die Fachleitungen im Schnitt mehr unterschiedliche Merkmale und beziehen sich häufiger auf Dimension II („Kognitive Aktivierung“), wohingegen die Referendar:innen häufiger Bezüge zu Dimension VII („Klassenführung“) aufweisen. Die Vernetzung der Merkmale führte sowohl in der Gesamtstichprobe als auch bei einer getrennten Betrachtung der Gruppen (FF2) zu jeweils vier Clustern. Die Merkmale innerhalb der Cluster überschneiden sich zwischen den Gruppen jedoch lediglich zu 33 % - 50 %. Die Modularität der Cluster fällt bei der Gesamtgruppe (0.12) geringer aus, als bei einer getrennten Analyse der Gruppen (Fachleitungen = 0.19, Referendar:innen = 0.16).

Für die Clusterbildung sind vor allem Merkmale von Interesse, die zu einer Vernetzung unterschiedlicher Cluster führen und damit die Trennschärfe verringern. Hierbei weisen beide Gruppen eine regelmäßige Nutzung des Merkmals „Auswahl und Einbindung von Fachinhalten“ (Dim. I) und der „Auswahl und Einbindung naturwissenschaftlicher Denk- und Arbeitsweisen“ (Dim. I) zur Inter-Cluster-Vernetzung auf. Darüber hinaus unterscheiden sich die Gruppen jedoch bei den Merkmalen, sodass Fachleitungen zusätzlich „Strukturierter Ablauf und Sequenzierung der Stunde“ (Dim. I), die „Förderung der Schüler:innenbeteiligung“ (Dim. V) und die „Motivierende Einbettung der Inhalte“ (Dim. I)

zur Vernetzung nutzen. Auf der Seite der Referendar:innen werden zusätzlich die Merkmale „Strukturierung und Fokussierung von Schlüsselaspekten“ (Dim. I) die „Konstruktive Einbindung von eigenen Ideen und Schüler:innenvorstellungen in den Unterricht“ (Dim. II) und die „Progression innerhalb der Stunde“ (Dim. I) verwendet. Die Inter-Cluster-Verbindungen sorgen zwar in beiden Gruppen für eine regelmäßige Vernetzung der Cluster, allerdings führt eine Ausnahme der Inter-Cluster-Verbindungen nicht zu einer Verbesserung der Modularität. Dies ist lediglich der Fall, wenn beide Gruppen als eine Gesamtstichprobe betrachtet werden, wobei ein deutlicher Anstieg der Modularität von 0.12 auf 0.21 auftritt. Unabhängig von der Betrachtung getrennter Gruppen oder einer separaten Analyse ergibt die Gegenüberstellung der Cluster mit der theoretischen Grundlage des SEP-Frameworks (FF 3) keine systematische Übereinstimmung. Durch die Clusteranalyse wurden lediglich vier Cluster berechnet, was eine Zusammenlegung der sieben Dimensionen des SEP-Frameworks notwendig macht. Die Merkmale innerhalb einer Dimension sind jedoch auf unterschiedliche Cluster verteilt.

Diskussion

Bei der offenen Unterrichtsqualitätsbeurteilung wurden einige wenige Merkmale als zentrale Vernetzungselemente genutzt. In beiden Gruppen umfasst die Inter-Cluster-Vernetzung häufig Merkmale aus Dimension I („Auswahl und Thematisierung von Inhalten“). Dies deutet grundsätzlich darauf hin, dass zumindest teilweise ein gemeinsames Verständnis von Unterrichtsqualität vorliegt. Häufig aufzufindende Abweichungen in der Operationalisierung der Unterrichtsqualität (Alp Christ et al., 2022) führen jedoch dazu, dass die Merkmale oftmals nach einer individuellen Logik vernetzt werden, was sich in den gruppenspezifischen Unterschieden der Cluster zeigt.

Die Gesprächsabschnitte blieben nicht systematisch innerhalb eines Clusters oder innerhalb einer Dimension der Unterrichtsqualität. Vielmehr bewegen sie sich über Bezüge zu einigen wenigen zentralen Merkmalen, häufig zwischen unterschiedlichen Bereichen der Unterrichtsqualität. Diesen zentralen Merkmalen wird somit eine sehr hohe Bedeutung zugewiesen und potenziell könnte deren Beurteilung die Beurteilung anderer Merkmale beeinflussen. Auffällig ist dabei, dass alle Teilnehmenden zwar die aus ihrer Sicht wichtigsten Merkmale benennen konnten, diese jedoch von den häufigsten Inter-Cluster-Verbindungen größtenteils abweichen.

Die individuellen und gruppenspezifischen Unterschiede könnten die Kommunikation in der Praxis erschweren, wenn die jeweiligen Vernetzungen nicht transparent sind. Dabei kommt hinzu, dass die induktiv gebildeten Cluster zwischen den beiden Gruppen nur eine teilweise Überschneidung aufweisen, was die Vernetzung der Merkmale ebenfalls schwerer nachvollziehbar machen kann. Zusätzlich weichen die induktiven Cluster von theoretischen Grundlagen zur Unterrichtsqualität ab. Dadurch wird die Notwendigkeit einer gemeinsamen Grundlage zur Unterrichtsqualitätsbeurteilung deutlich. Grundsätzlich bietet die Abweichung zwischen induktiver Clusterung von Unterrichtsqualitätsmerkmalen in der Praxis und theoretischen Grundlagen einen Anlass, um die Strukturen der Unterrichtsqualität zu überdenken und auch den Ergänzungsbedarf zentraler Systematiken, wie den drei Basisdimensionen, (Praetorius et al., 2020b) aus einer zusätzlichen Perspektive zu betrachten. Die Relevanz von Theorien für die Praxis bleibt bestehen. Es könnte jedoch zielführend sein, bestehende Strukturen in der Praxis zu analysieren und zu berücksichtigen, um die Vernetzung von Theorie und Praxis von beiden Seiten aus produktiv zu gestalten und damit den Pluralismus an Unterrichtsqualitätsverständnissen zielführend zu reduzieren.

Literatur

- Alp Christ, A., Capon-Sieber, V., Grob, U. & Praetorius, A.-K. (2022). Learning processes and their mediating role between teaching quality and student achievement: A systematic review. *Studies in Educational Evaluation*, 75, 101209. <https://doi.org/10.1016/j.stueduc.2022.101209>
- Heinitz, B. & Nehring, A. (2023). Instructional quality in science teacher education: comparing evaluations by chemistry pre-service teachers and their advisors. *International Journal of Science Education*. <https://doi.org/10.1080/09500693.2023.2213382>
- Klieme, E., Schümer, G. & Knoll, S. (2001). Mathematikunterricht in der Sekundarstufe I. „Aufgabenkultur“ und Unterrichtsgestaltung. In E. Klieme & J. Baumert (Hrsg.), *TIMSS-Impulse für Schule und Unterricht. Forschungsbefunde, Reforminitiativen, Praxisberichte und Video-Dokumente* (S. 43–57). Bonn: Bundesministerium für Bildung und Forschung.
- Praetorius, A. K., Herrmann, C., Gerlach, E., Zülsdorf-Kersting, M., Heinitz, B. & Nehring, A. (2020a). Teaching quality in different subject matters in German-speaking countries—Inbetween genericness and subject-specificity. *Unterrichtswissenschaft*, 48(3), 409–446. <https://doi.org/10.1007/s42010-020-00082-8>
- Praetorius, A.-K., Rogh, W., & Kleickmann, T. (2020b). Blinde Flecken des Modells der drei Basisdimensionen von Unterrichtsqualität? Das Modell im Spiegel einer internationalen Synthese von Merkmalen der Unterrichtsqualität. *Unterrichtswissenschaft*, 48(3). <https://doi.org/10.1007/s42010-020-00072-w>
- Rothland, M. (2024). Auf Sand gebaut? Über den Versuch, die Basisdimensionen der Unterrichtsqualität in erziehungswissenschaftlicher „Theorie“ zu verankern. *Zeitschrift Für Bildungsforschung*, 14(1), 41–53. <https://doi.org/10.1007/s35834-024-00411-2>
- Weber K. E. & Czerwenka K. (2021). Anschlussfähigkeit und Kooperation der ersten und zweiten Phase der Lehrkräftebildung. In J. Peitz J. & M. Harring M. (Hrsg.), *Das Referendariat. Ein systematischer Blick auf den schulpraktischen Vorbereitungsdienst* (S. 255 – 264). Waxmann.