

Benjamin Heinitz<sup>1</sup>  
Andreas Nehring<sup>1</sup>

<sup>1</sup>Leibniz Universität Hannover

## **Beurteilung von Unterrichtsqualität im Referendariat**

### **Projekthintergrund**

Die Bewertung von Unterrichtsqualität nimmt in der zweiten Phase der Lehramtsausbildung eine zentrale Rolle ein. Im Verlauf des Referendariats finden regelmäßige Unterrichtsbesuche durch Fach- bzw. Seminarleiter\*innen statt sowie ein abschließender Prüfungsunterricht, in denen der Unterricht der Referendar\*innen bewertet wird. Diese Unterrichtsbesuche sind außerdem an Reflexionsgespräche gekoppelt, in denen die Beobachtungen, Interpretationen und Bewertungen als Beratungsgrundlage für die weitere Ausbildung genutzt werden.

Die Wahrnehmung von Unterrichtsqualität wurde in den letzten Jahren bereits in vielen Aspekten detailliert untersucht: So konnte bei einer kritischen Betrachtung der Beobachterperspektive festgestellt werden, dass ein Beurteilereffekt zu einer Varianz in der Bewertung führen kann (Praetorius, 2013). Bei einer Untersuchung der Wahrnehmung von Unterrichtsqualität aus unterschiedlichen Perspektiven (Lehrer\*innen, Schüler\*innen, Beobachter\*innen), konnte festgestellt werden, dass die jeweiligen Perspektiven einen Einfluss auf die Bewertung haben können (Fauth, Decristan, Rieser, Klieme, & Büttner, 2014). Aber auch bei einer Untersuchung der Unterrichtsqualität durch unterschiedliche Instrumente konnte eine Varianz in der Bewertung festgestellt werden (Brunner, 2018). Die Gruppe der Fach- und Seminarleiter\*innen als Beurteiler von Unterrichtsqualität sind jedoch weitgehend unerforscht. Auch wenn ein Zusammenspiel der individuellen Faktoren einen Unterschied in der Wahrnehmung der Unterrichtsqualität bereits anbahnt, stellt sich doch die Frage, inwiefern dies zu einer abweichenden Bewertung einer Unterrichtsstunde führt.

### **Zielstellung**

Vor diesem Hintergrund bestand das Ziel darin eine direkte Gegenüberstellung der Bewertung einer Unterrichtsstunde durch verschiedene Fach- und Seminarleiter\*innen zu erstellen. Die konkreten Fragestellungen waren hierbei:

Welche Kriterien für Unterrichtsqualität werden zur Bewertung verwendet und wie werden diese eingeschätzt?

Wie homogen ist die Gruppe der Fach- und Seminarleiter\*innen in der Bewertung derselben Unterrichtsstunde?

### **Methodisches Vorgehen**

Um den Fach- und Seminarleiter\*innen eine gemeinsame Basis der Bewertung zu bieten, wurde im Vorfeld ein Video aufgezeichnet, das eine Referendarin in einer 45-minütigen Chemiestunde zeigt. Diese Stunde wurde dann in 17 leitfadengestützten Einzelinterviews als Grundlage für die Bewertung einer typischen Chemiestunde verwendet, wobei der Unterrichtsentwurf den Interviewten im Vorfeld zur Verfügung gestellt wurde. Die Interviews wurden vollständig aufgezeichnet, transkribiert, in Kodiereinheiten unterteilt und anschließend kodiert. Da die Nutzung der Kriterien freigestellt war, wurde für die Kodierung eine möglichst umfangreiche Systematisierung von Unterrichtsqualität benötigt, welche in Form der naturwissenschaftsdidaktischen Perspektivierungen (Heinitz & Nehring, 2020)

vorlag. Diese 50 Perspektivierungen enthalten sowohl generische, als auch fachspezifische und hybridisierte Kriterien und stehen darüber hinaus in einem Bezugsrahmen zum generischen Syntheseframework (Praetorius, Herrmann, Gerlach, Zülsdorf-Kersting, Heinitz & Nehring, 2020), wodurch Kriterien für Unterrichtsqualität in diesem System vergleichbar verortet werden können. Neben der Zuteilung der Kodiereinheiten zu den 50 naturwissenschaftsdidaktischen Perspektivierungen wurde auch die Qualität der Wahrnehmung durch die Fach- und Seminarleiter\*innen in vier Stufen erfasst und ebenso die Wertung des jeweiligen Kriteriums für Unterrichtsqualität durch die Fach- und Seminarleiter\*innen in fünf Abstufungen.

Die vier Stufen der Wahrnehmungsqualität umfassen dabei:

1. Die einfache Nennung einer Perspektivierung als Kriterium für Unterrichtsqualität.
2. Die Nennung einer Perspektivierung mit einem Bezug zur angesehenen Stunde.
3. Die Nennung einer Perspektivierung mit einem Vorschlag zur Verbesserung der angesehenen Stunde.
4. Die Nennung und Vernetzung einer Perspektivierung mit anderen Kriterien der Unterrichtsqualität.

Die fünf Stufen der Wertung einer Perspektivierung durch die Fach- und Seminarleiter\*innen beziehen sich jeweils auf die Gesamtheit aller, in einer Perspektivierung verorteten, Aussagen. Dabei werden die einzeln kodierten Aussagen der Interviewten untersucht und mit einer Wertung (positiv, neutral, negativ) versehen. Anschließend wurden alle Aussagen, die einer Perspektivierung zugeordnet wurden gesammelt und die Gesamtwertung in einem fünfstufigen Spektrum von „vollständig positiv“ bis „vollständig negativ“ verortet. Zusätzlich wurde die Benotung der Unterrichtsstunde vor zwei potentiellen Hintergründen erfasst: 1. Die Benotung der Stunde in einem Unterrichtsbesuch im 12. Monat des Referendariats (Abb. 1 blau). 2. Die Benotung der Stunde als Prüfungsunterricht am Ende des Referendariats (Abb. 1 orange).

### Ergebnisse

Die Bewertungen der Unterrichtsstunde durch die Fach- und Seminarleiter\*innen fielen insgesamt sehr unterschiedlich aus. Hierbei unterschieden sie sich sowohl in den ausgewählten

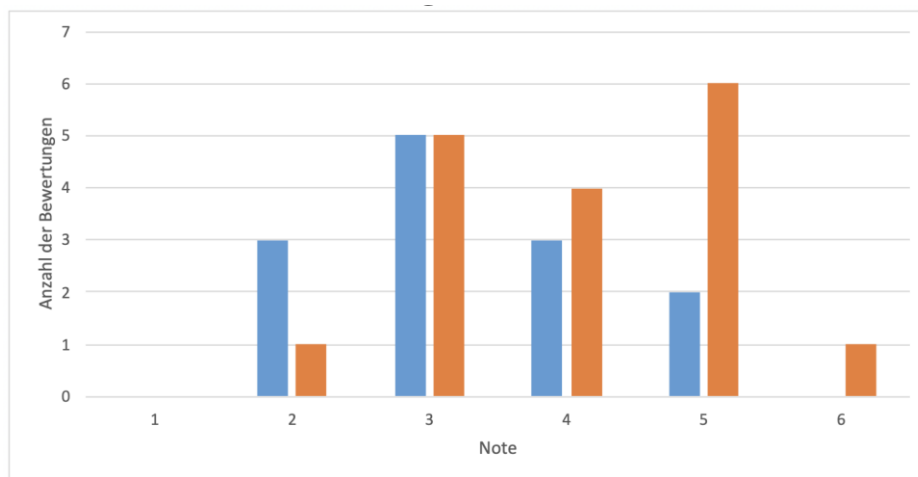


Abbildung 1: Benotung der Unterrichtsstunde durch die Fach- und Seminarleiter\*innen

Qualitätskriterien, als auch in der Wertung derselben Kriterien. Diese heterogenen Wahrnehmungen und Wertungen der Unterrichtsqualität bilden sich dann auch in der Benotung der Unterrichtsstunde (Abb. 1) ab. Für den Prüfungsunterricht würde aus dieser heterogenen Benotung folgen, dass dieselbe Referendarin mit derselben Stunde in 10 Fällen bestanden hätte und in 7 Fällen durchgefallen wäre. Bei einer Betrachtung der einfachen Nutzung von Kriterien zur Bewertung der Unterrichtsstunde lassen sich 3 Typen von Perspektivierungen herausstellen: 1. Perspektivierungen die von (fast) allen Interviewten für die Benotung herangezogen wurden. 2. Perspektivierungen die nur von einem Teil der Interviewten genutzt wurden. 3. Perspektivierungen die nicht, oder nur selten zur Bewertung genutzt wurden. Vor allem der zweite Fall wirft hierbei die Frage auf, ob ein Teil der Fach- und Seminarleiter\*innen die jeweilige Perspektivierung nicht wahrgenommen, oder lediglich nicht als relevantes Kriterium eingestuft haben. Hierfür bietet die „Aktivierung von Vorwissen“ ein gutes Beispiel, da sie von knapp der Hälfte (9/17) der Interviewten genutzt wurde und in 8 Fällen sogar ein Verbesserungsvorschlag für die Stunde angeboten wurde, wohingegen die andere Hälfte diese Perspektivierung nicht nutzte. Weiterhin können Unterschiede in der Bewertung der Kriterien einen guten Ansatz für eine detaillierte Untersuchung bieten. Auch hierbei konnten, selbst bei Kriterien die von allen Interviewten genutzt wurden, starke Unterschiede herausgestellt werden. Im Fall der „Motivierenden Einbettung von Inhalten“ wird beispielsweise zwischen einer „vollständig positiven“ und einer „vollständig negativen“ Bewertung jede mögliche Abstufung durch mindestens einen der Fach- und Seminarleiter\*innen erreicht. Weiterhin können auch bei Perspektivierungen, die im Gesamtbild relativ konsistent bewertet wurden, z.B. die „Auswahl und Einbindung von Fachinhalten“, widersprüchliche Aussagen gefunden werden.

### **Diskussion**

Die Auswertung und Gegenüberstellung der Interviews und der Bewertungen einer Unterrichtsstunde verdeutlicht, dass die Gruppe der Fach- und Seminarleiter\*innen trotz einer theoretisch vergleichbaren Expertise vergleichsweise heterogen ist. Dies kann vor allem durch die individuellen Hintergründe der Interviewten und fehlender einheitlicher Voraussetzungen für die Besetzung einer Fach- bzw. Seminarleiter\*innenstelle liegen, zumal die Interviewten teilweise in unterschiedlichen Bundesländern tätig sind. Hierbei bleibt die Frage offen, inwiefern dies für die Lehramtsausbildung eine geeignete Basis für die Ausbildung von Referendar\*innen darstellt und ob diese für eine vergleichbare Qualifikation stärker vereinheitlicht werden sollte. Schließlich kann die Notenvergabe die Bewerbungs- und Karrierechancen für den Schuldienst stark beeinflussen.

Auch können die Unterschiede in Wahrnehmung und Bewertung, die im Verlauf dieser Untersuchung herausgestellt wurden, gut für eine weiterführende Untersuchung bzw. eine Intervention genutzt werden, um die Bewertung von Unterrichtsstunden im Referendariat gezielt vergleichbarer zu gestalten.

### **Danksagung:**

Wir danken den 17 Fach- und Seminarleiter\*innen für Ihre Teilnahme an den Interviews und die tiefgehenden Einblicke in die zweite Phase der Lehramtsausbildung die uns dadurch geboten wurde.

Wir danken außerdem Jasmin Meyer und Jos Oldag für die Durchführung der Interviews.

**Literatur**

- Brunner, E. Qualität von Mathematikunterricht: Eine Frage der Perspektive. *J Math Didakt* 39, 257–284 (2018). <https://doi.org/10.1007/s13138-017-0122-z>
- Fauth, Benjamin, Decristan, Jasmin, Rieser, Svenja, Klieme, Eckhard und Büttner, Gerhard (2014) Grundschulunterricht aus Schüler-, Lehrer- und Beobachterperspektive: Zusammenhänge und Vorhersage von Lernerfolg, *Zeitschrift für Pädagogische Psychologie*, 28 (3), 2014, 127–137
- Heinitz, B., Nehring, A. Kriterien naturwissenschaftsdidaktischer Unterrichtsqualität – ein systematisches Review videobasierter Unterrichtsforschung. *Unterrichtswiss* 48, 319–360 (2020). <https://doi.org/10.1007/s42010-020-00074-8>
- Praetorius, A. (2013) Einschätzung von Unterrichtsqualität durch externe Beobachterinnen und Beobachter. Eine kritische Betrachtung der aktuellen Vorgehensweise in der Schulpraxis, *Beiträge zur Lehrerinnen- und Lehrerbildung* 31 (2013) 2, S. 174-185
- Praetorius, A., Herrmann, C., Gerlach, E., Zülsdorf-Kersting, M., Heinitz, B., & Nehring, A., Unterrichtsqualität in den Fachdidaktiken im deutschsprachigen Raum – zwischen Generik und Fachspezifik. *Unterrichtswiss* 48, 409–446 (2020). <https://doi.org/10.1007/s42010-020-00082-8>