

Geschichtsdidaktik empirisch 17: Translation II

## **«Aber was hier bei dieser Schülerin komplett gefehlt hat, ist [...] dieses historische Denken.» Geschichtslehrkräfte beurteilen Schülerleistungen – erste Ergebnisse einer qualitativ-empirischen Studie**

*Inga Kahlcke*<sup>\*</sup>

Die Beurteilung schriftlicher Schülerleistungen zählt zu den unumgänglichen Aufgaben des (Geschichts-)Lehrerberufs. Nicht zuletzt seit dem Einzug der Kompetenzorientierung und der damit einhergehenden Output-Orientierung in die geschichtsdidaktische Diskussion stellt sich das Problem der Messung bzw. der Messbarkeit historischer Kompetenzen in Forschung und schulischer Praxis mit besonderer Dringlichkeit (siehe etwa Körber 2008). Zwar existieren mittlerweile einige pragmatisch ausgerichtete Handreichungen zur kompetenzorientierten Leistungsbeurteilung im Geschichtsunterricht (Adamski 2014; Kühberger 2014). Für die empirische Forschung kann aber insbesondere bezüglich der alltäglichen Leistungsbeurteilungspraxis von Geschichtslehrkräften weiterhin nur das «Fehlen eines vertretbaren Forschungsstandes» (Borries 2007, 654) konstatiert werden. Erste Ergebnisse aus anderen Forschungszusammenhängen zeichnen diesbezüglich ein eher düsteres Bild: So kommen Bernd Schönemann, Holger Thünemann und Meik Zülsdorf-Kersting (2010, 98) in ihrer Untersuchung von Abiturklausuren zu dem Schluss, dass einerseits einige Leistungen, die von den Autoren als deutlich defizitär angesehen wurden, von den schulischen Beurteilern mit (sehr) guten Noten bedacht wurden, während andererseits in Einzelfällen Ansätze elaborierten historischen Denkens unterbewertet wurden. Dies lässt die Vermutung zu, dass zwischen geschichtsdidaktischen Ansprüchen und schulischer Praxis der Leistungsbeurteilung eine Kluft besteht, die erstens genauer beschrieben, zweitens nach Möglichkeit erklärt und drittens durch gezielte Maßnahmen verringert werden sollte.

Im Vortrag wird ein Dissertationsprojekt vorgestellt, das diese drei Ziele verfolgt. Es nimmt die Leistungsbeurteilung als Teil der Professionalität von Geschichtslehrkräften in den Blick. Die Studie verortet sich an der Schnittstelle von erziehungswissenschaftlicher Professionsforschung und geschichtsdidaktischer Kompetenzforschung und geht qualitativ-deskriptiv vor. Anders als in der pädagogisch-psychologischen Literatur, die schon seit Jahrzehnten die «Fragwürdigkeit der Zensurengebung» (Ingenkamp 1971) konstatiert, geht es dabei nicht darum, Urteile von Lehrkräften anhand testtheoretischer Gütekriterien zu überprüfen. Vielmehr werden in Anlehnung an kognitionspsychologische (Krolak-Schwerdt, Böhmer & Gräsel 2012; Lissmann 1987) und ethnographische Studien (Kalthoff 1996; Lüders 2001; Terhart 2000; Zaborowski, Meier & Breidenstein 2011) individuelle Leistungsbeurteilungspraktiken in

---

<sup>\*</sup> Didaktik der Geschichte, Georg-August-Universität Göttingen

geschichtsdidaktischer Perspektive erforscht: Untersucht wird, wie Lehrkräfte bei der schriftlichen Leistungsbeurteilung im Geschichtsunterricht vorgehen, auf welchen Wissensbeständen und auf welchen berufsbezogenen Überzeugungen (Reusser & Pauli 2014; vgl. auch Wineburg & Wilson 2001) dieses Vorgehen beruht und inwiefern es mit der Berufserfahrung von Geschichtslehrkräften zusammenhängt. Auf diese Weise sollen Ansatzpunkte für die Aus- und Fortbildung von Geschichtslehrkräften im Bereich der Leistungsbeurteilung identifiziert werden. Im Projekt werden gymnasiale Geschichtslehrkräfte von zehnten Klassen in einer nach Berufserfahrung (Referendare, Berufsanfänger und erfahrene Lehrkräfte) geschichteten Stichprobe ( $N = 15$ ) befragt.

Die Datenerhebung folgt dem Prinzip der Triangulation: Es wird die Methode des Lauten Denkens (Konrad 2010) mit problemzentrierten Interviews (Witzel 1985) kombiniert. Das Laute Denken dient der Erhebung der Leistungsbeurteilungspraxis. Dabei korrigieren die Lehrkräfte erstens Klassenarbeiten, die sie in ihren eigenen zehnten Klassen durchgeführt haben, und zweitens eine anonymisierte Klassenarbeit einer neunten Klasse, die allen Lehrkräften gleichermaßen vorgelegt wird. Die Beobachtung der Korrektur einer eigenen Klassenarbeit ermöglicht einen direkten Einblick in die tatsächliche Leistungsbeurteilungspraxis der Lehrkräfte, während die Korrektur der Fremdklassenarbeit den Vergleich der unterschiedlichen Bewertungsschwerpunkte der Lehrkräfte ermöglicht. Die problemzentrierten Interviews dienen insbesondere der Erhebung der berufsbezogenen Überzeugungen. Die Datenerhebung erfolgt seit November 2016 und wird voraussichtlich im Sommer 2017 abgeschlossen sein.

Die Auswertung des Materials erfolgt anhand der dokumentarischen Methode (Bohnsack 2007; Nohl 2012). Dabei werden drei verschiedene Perspektiven an das Material angelegt. Die erste Perspektive entspricht in der Nomenklatur der dokumentarischen Methode der sinngenetischen Typenbildung, während die zweite und dritte Perspektive der soziogenetischen Typenbildung entsprechen (Nohl 2012, 51–57).

Unter der ersten Perspektive werden zunächst diejenigen Kriterien der Leistungsbeurteilung identifiziert, die in fast allen Laut-Denk-Protokollen auftauchen, mit denen die Lehrkräfte aber erkennbar unterschiedlich umgehen. Als solche Kriterien haben sich bisher Quellenkritik, Sprache, Sachwissen und (Wert-)urteile erwiesen. Die unterschiedlichen Umgangsweisen der Lehrkräfte mit diesen wiederholt vorkommenden Aspekten werden in der dokumentarischen Methode als verschiedene Orientierungsrahmen bezeichnet (Nohl 2012, 49).

Unter der zweiten und dritten Perspektive wird untersucht, in welchem sozialen Zusammenhang diese unterschiedlichen Orientierungsrahmen stehen (Nohl 2012, 52). In der zweiten Perspektive werden die berufsbezogenen Überzeugungen als sozialer Zusammenhang fokussiert. Wie das Zitat im Titel des geplanten Vortrags zeigt, spielen etwa Überzeugungen dazu, was historisches Denken ausmacht, in der Leistungsbeurteilungspraxis eine wichtige Rolle. Daher werden in diesem Schritt die Ergebnisse der Laut-Denk-Protokolle mit den Interviewdaten in Beziehung gesetzt. Es wird etwa untersucht, inwiefern unterschiedliche Ziele von Geschichtslehrkräften (als Teil der berufsbezogenen Überzeugungen) mit Unterschieden

«Aber was hier bei dieser Schülerin komplett gefehlt hat, ist [...] dieses historische Denken.»

in der Leistungsbeurteilungspraxis einhergehen. Dabei haben sich bereits einige der unter der ersten Perspektive herausgearbeiteten Orientierungsrahmen als konsistent mit unterschiedlichen Überzeugungen der Lehrkräfte erwiesen. Zugleich haben sich jedoch auch beträchtliche Diskrepanzen zwischen den Überzeugungen der Lehrkräfte und ihrer Leistungsbeurteilungspraxis gezeigt: So verwenden mitunter auch Lehrkräfte, deren explizites Ziel die Vermittlung des Konstruktcharakters von Geschichte ist, die Menge des gelernten Faktenwissens als wichtigstes Kriterium. Daher wird das Material auch danach analysiert, wie diese Diskrepanzen erklärt werden können. Bisher konnten zeitökonomische Gründe, skeptische Haltungen gegenüber der Messbarkeit gerade elaborierter historischer Kompetenzen sowie Überzeugungen zur Entwicklung des historischen Denkens als Ursachen identifiziert werden; es wird erwartet, dass aus den weiteren Interviews noch andere Gründe rekonstruiert werden können.

In der dritten Perspektive wird untersucht, inwiefern die unterschiedlichen Orientierungsrahmen mit der Berufserfahrung der Lehrkräfte in Zusammenhang stehen. Dafür werden die drei untersuchten Erfahrungsgruppen miteinander kontrastiert. Die Ergebnisse aus diesem Schritt sollen Hinweise darauf liefern, wie sich die Expertise von Geschichtslehrkräften im Bereich der Leistungsbeurteilung entwickelt und wie dieser Prozess unterstützt werden kann.

Im Vortrag wird nach einem Abriss über die Probleme der kompetenzorientierten Leistungsbeurteilung im Fach Geschichte die Konzeption des Forschungsprojekts vorgestellt. Anschließend werden anhand von beispielhaften Analysen einiger Ausschnitte aus den LautDenk-Protokollen und den Interviews erste Ergebnisse bezüglich der drei oben skizzierten Perspektiven präsentiert. Dabei wird auch das konkrete methodische Vorgehen bei der Auswertung des Materials verdeutlicht. Den Schluss bilden ein Ausblick auf die weiteren Forschungsfragen sowie erste pragmatische Forderungen, die sich aus den Ergebnissen ableiten lassen.

## Literatur

- Adamski, Peter (2014). Historisches Lernen Diagnostizieren. Lernvoraussetzungen – Lernprozesse – Lernleistungen. Schwalbach/Ts: Wochenschau.
- Bohnsack, Ralf (2007). *Rekonstruktive Sozialforschung: Einführung in qualitative Methoden* (6. Aufl.). Opladen: Budrich.
- Borries, Bodo von (2007). Empirie: Ergebnisse messen (Lerndiagnose im Fach Geschichte). In Andreas Körber, Waltraud Schreiber & Alexander Schöner (Hrsg.), *Kompetenzen historischen Denkens. Ein Strukturmodell als Beitrag zur Kompetenzorientierung in der Geschichtsdidaktik* (S. 653–673). Neuried: Ars una.
- Ingenkamp, Karlheinz (Hrsg.) (1971). *Die Fragwürdigkeit der Zensurengebung: Texte und Untersuchungsberichte*. Weinheim: Beltz.
- Kalthoff, Herbert (1996). Das Zensurenpanoptikum. Eine ethnographische Studie zur schulischen Bewertungspraxis. *Zeitschrift für Soziologie*, 25(2), 106–124.

- Konrad, Klaus (2010). Lautes Denken. In Günter Mey & Katja Mruck (Hrsg.), *Handbuch Qualitative Forschung in der Psychologie* (S. 476–490). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Körper, Andreas (2008). Sind Kompetenzen historischen Denkens messbar? In Volker Frederking (Hrsg.), *Schwer messbare Kompetenzen: Herausforderungen für die empirische Fachdidaktik* (S. 65–84). Baltmannsweiler: Schneider.
- Krolak-Schwerdt, Sabine, Böhmer, Matthias, & Gräsel, Cornelia (2012). Leistungsbeurteilung von Schulkindern. Welche Rolle spielen Ziele und Expertise der Lehrkraft? *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 44(3), 111–122.
- Kühberger, Christoph (2014). *Leistungsfeststellung im Geschichtsunterricht: Diagnose – Bewertung – Beurteilung*. Schwalbach/Ts: Wochenschau.
- Lissmann, Urban (1987). Lehrgedanken zur Schülerbeurteilung. Dimensionalität und Struktur. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 19(3), 266–284.
- Lüders, Manfred (2001). Probleme von Lehrerinnen und Lehrern mit der Beurteilung von Schülerleistungen. *Zeitschrift für Erziehungswissenschaft*, 4(3), 457–474.
- Nohl, Arnd-Michael (2012). *Interview und dokumentarische Methode: Anleitungen für die Forschungspraxis* (4. Aufl.). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Reusser, Kurt, & Pauli, Christine (2014). Berufsbezogene Überzeugungen von Lehrerinnen und Lehrern. In Ewald Terhart, Hedda Bennewitz & Martin Rothland (Hrsg.), *Handbuch der Forschung zum Lehrerberuf* (2. Aufl., S. 642–661). Münster: Waxmann.
- Schönemann, Bernd, Thünemann, Holger, & Zülsdorf-Kersting, Meik (2010). *Was können Abiturienten? Zugleich ein Beitrag zur Debatte über Kompetenzen und Standards im Fach Geschichte*. Berlin Lit.
- Terhart, Ewald (2000). Schüler beurteilen – Zensuren geben. Wie Lehrerinnen und Lehrer mit einem leidigen, aber unausweichlichen Element ihres Berufsalltags umgehen. In Silvia-Iris Beutel & Witlof Vollstädt (Hrsg.), *Leistung ermitteln und bewerten* (S. 39–50). Hamburg: Bergmann und Helbig.
- Wineburg, Sam, & Wilson, Suzanne (2001). Wrinkles in Time and Place. Using Performance Assessments to Understand the Knowledge of History Teachers. In Sam Wineburg (Hrsg.), *Historical Thinking and Other Unnatural Acts: Charting the Future of Teaching the Past* (S. 173–214). Philadelphia: Temple University Press.
- Witzel, Andreas (1985). Das problemzentrierte Interview. In Gerd Jüttemann (Hrsg.), *Qualitative Forschung in der Psychologie* (S. 227–255). Weinheim: Beltz.
- Zaborowski, Katrin Ulrike, Meier, Michael, & Breidenstein, Michael (2011). *Leistungsbewertung und Unterricht. Ethnographische Studien zur Bewertungspraxis in Gymnasium und Sekundarschule*. Wiesbaden: VS Verlag für Sozialwissenschaften.