

Lukas Mientus¹
Anna Nowak¹
Peter Wulff²
Andreas Borowski¹

¹Universität Potsdam
²PH Heidelberg

Unterrichtspraxis reflektieren – automatisiert Rückmeldung geben

Reflexion wird oft als Schlüsselkategorie in der professionellen Entwicklung von Lehrkräften diskutiert (Korthagen & Kessels, 1999; Combe & Colbe, 2008; Christof et al., 2018; Wyss & Mahler, 2021). Unterrichtsreflexion kann einen nachhaltigen und intrinsischen Professionalisierungsprozess anregen (Roters, 2012), so dass sie unter anderem unabdingbar für die Entwicklung einer Lehrpersönlichkeit scheint (Rothland, 2021). Reflexion kann definiert werden als ein strukturierter Prozess im konkreten Unterrichtskontext mit dem Ziel der persönlichen Weiterentwicklung (nach Mientus, 2023 in Anlehnung an von Aufschnaiter et al., 2019) und kann somit nicht nur durch inhaltliches, sondern auch durch strukturelles Feedback gefördert werden (Poldner et al., 2014). Ein lernwirksames Feedback nach Hattie und Timperley (2007), das sowohl eine Rückmeldung zum derzeitigen Stand der professionellen Entwicklung als auch Zukunftsperspektiven aufweisen sollte, stellt sich hierbei jedoch als ressourcenintensiv dar, weshalb im Bereich der Reflective Writing Analytics Abhilfe gefunden werden kann etwa durch die Automatisierung solcher ressourcenintensiven Prozesse (Lai & Calandra, 2007; Buckingham Shum et al., 2017; Ullman, 2019; Wulff et al., 2022). Inwieweit ein konkretes Feedbackkonzept, welches strukturelles Feedback zu schriftlichen Reflexionen computerbasiert bereitstellt, entgegen einer Vergleichsgruppe gewinnbringend ist, soll in der vorliegenden Studie erprobt werden.

Forschungsanliegen im Kohortenvergleich

Im Feedbackkonzept *ReFeed* (Mientus et al., 2019) werden basierend auf Algorithmen aus Wulff et al. (2022) automatisch Rückmeldungen zur Argumentationsstruktur nach dem Rahmenmodell für Reflexion nach Nowak et al. (2019) gegeben. Nach dem Modell besteht ein höherwertiger Reflexionsprozess darin, dass in einem ersten Schritt Beobachtungen beschrieben werden. Auf Grundlage des individuellen Professionswissens können dann Bewertungen vorgenommen und Alternativen zur Verbesserung des Unterrichts sowie Konsequenzen zur eigenen professionellen Entwicklung formuliert werden (Nowak, 2023). Um die Wirksamkeit des Feedbacks nach Hattie und Timperley (2007) zu validieren, sollen $N = 132$ schriftliche Selbstreflexionen von $N = 22$ Physik-Lehramtsstudierenden aus drei Semestern (SoSe 2017 bis SoSe 2018) als Vergleichsgruppe mit einer Interventionsgruppe, welche mit *ReFeed* begleitet wurde, verglichen werden. Die Interventionsgruppe umfasst $N = 102$ Selbstreflexionen von $N = 34$ Studierenden aus fünf Semestern (WiSe 2020|21 bis WiSe 2022|24) (3 Reflexionen p.P.). Die Vergleichbarkeit beider Kohorten ist unter Mientus et al. (eingereicht) diskutiert und wird in diesem Beitrag angenommen. Entscheidend ist jedoch, dass die Vergleichsgruppe ohne explizites, individuelles Feedback Reflexionen verfasste, wohingegen die Interventionsgruppe im Sinne eines *Feed Up* (nach Hattie & Timperley die Klärung der Erwartungshaltung an eine Reflexion) mit dem Rahmenmodell für Reflexion nach Nowak et al. (2019) als Prompt auf die Reflexionspraxis vorbereitet wurde.

Automatisierte Klassifikationsmethoden

Für die Analyse der schriftlichen Reflexionen beider Kohorten wurden automatische Klassifikations- und Modellierungsalgorithmen aus Wulff et al. (2022) und Mientus et al. (2023) verwendet. Der Klassifikationsalgorithmus ordnet jedem Satz eines Reflexionstextes ein Element des Rahmenmodells nach Nowak et al. (2019) zu. Der prozentuale Anteil der Beschreibung wird hier als deskriptiver Anteil der Reflexion bezeichnet und die Summe der Anteile an Bewertungen, Alternativen und Konsequenzen als diskursiver Anteil. Die Verteilung beider Anteile kann als Indikator für Reflexionsqualität angesehen werden, wobei novizenhafte Texte eher deskriptiv ausfallen (Fund et al., 2002). Da auch die Positionierung der Elemente im Text ein Qualitätsindikator darstellen kann, repräsentiert der Modellierungsalgorithmus einen *Level of Structure*, welcher in Extremgruppen gelungen, moderat oder schwach strukturierte Texte Reflexionen höherer, mittlerer bzw. geringerer Qualität abbildet (Mientus et al., 2023).

Zusammenfassung der Ergebnisse

In beiden Kohorten und über alle Reflexionen hinweg liegt erwartungskonform der deskriptive Anteil der Reflexionen bei weit über 50% (siehe Abbildung 1). Auch sind keine signifikanten Änderungen der Verteilung beider Anteile im Verlauf des Praxissemesters feststellbar (siehe Abb. 1). Gleichzeitig startet die Interventionsgruppe bereits zur ersten schriftlichen Reflexion (also noch vor dem ersten Feedback) auf qualitativ höherem Niveau. Der Anteil diskursiver Elemente ist kontinuierlich höher als in der Vergleichskohorte.

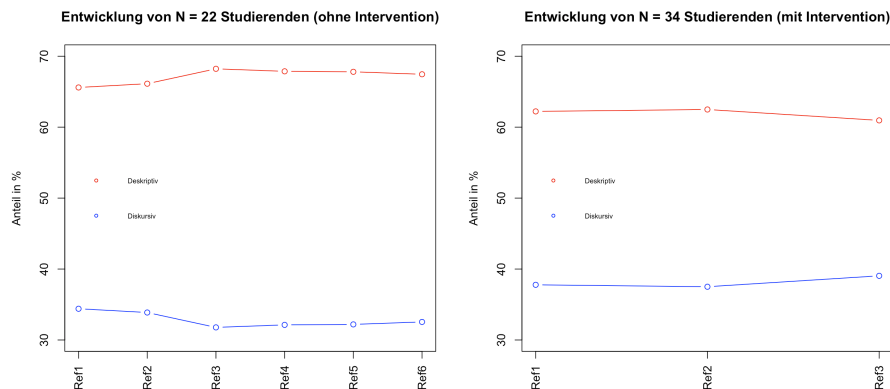


Abb. 1: Verlauf deskriptiver (rot) und diskursiver (blau) Anteile der Reflexionen

Betrachtet man den Verlauf der Extremgruppenanalyse des *Level of Structure* (Abb. 2) scheinen sich Personen, welche zu Beginn des Praxissemesters gelungen strukturieren, im Verlauf des Praxissemesters den anderen beiden Extremgruppen anzunähern. Diese Beobachtung trifft auf beide Kohorten zu. Moderat strukturierende Personen bleiben auf konstantem Niveau, wohingegen schwache Reflexionstexte schwach strukturiert bleiben. Nahezu keine Entwicklung (egal welcher Extremgruppe und Kohorte) wird hierbei signifikant. Einzige negative Entwicklung des *Level of Structure* nach dem ersten Feedback

aus der Kohorte mit Intervention zwischen dem ersten und dem zweiten Reflexionstext stellt mit $p = .003$ eine signifikante Veränderung dar.

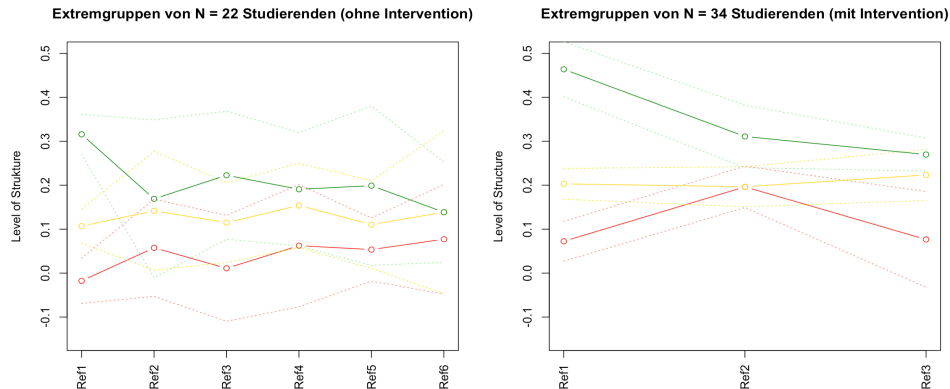


Abb. 2: Verlauf der gelungen (grün), moderat (gelb) und schwach (rot) strukturierten Reflexionen im Verlauf des Praxissemesters ohne und mit Intervention

Lessons Learned

Ohne Intervention scheinen sich die Strukturen der Extremgruppen einander anzunähern. Diese Beobachtung kann potenziell auf einen Messeffekt zurückzuführen sein. So beobachtete Nowak (2023) für dieselbe Stichprobe, dass die Länge der Reflexionstexte sinkt und gleichzeitig die Qualität konstant bleibt. Mientus et al. (2023) nahmen bei der Betrachtung des Level of Structure an, ein Qualitätsmaß identifiziert zu haben, welches von der Textlänge unabhängig war.

Bei der Kohorte mit Intervention verbleiben moderat und schwach strukturierende Personen auf konstantem Niveau. Das Rahmenmodell für Reflexion (als Feed Up) scheint wirksam, da im Kohortenvergleich mit zugehörigem Prompt weniger deskriptive Texte formuliert wurden. Gleichzeitig zeigte sich die Intervention (nach Hattie und Timperley als Feed Back + Feed Forward) als nicht eindeutig wirksam.

Diese Studie berücksichtigt das erteilte Feedback und vernachlässigt hierbei die generelle kognitive Belastung der Studierenden, welche im Verlauf des Praxissemesters zunehmen kann (z.B. Jantowski et al., 2010). Gleichzeitig zeigen Studien, dass eine emotionale Belastung sinkt (z.B. Römer et al., 2018), sodass potenziell das Bestreben, einen guten Reflexionstext zu verfassen abnimmt. Mit dem Feedbackkonzept kann zwar dem Wunsch nach zeitnahe, individuellem und gehaltvollem Feedback gerecht werden, messbare Effekte konnten jedoch nur bedingt identifiziert werden. Für einen entscheidenden Effekt, um höhere Qualität in schriftlichen Selbstreflexionen zu erreichen scheint eine Sensibilisierung für ein/das Rahmenmodell für Reflexion vor Beginn der Unterrichtspraxis zu genügen.

Literatur

- Buckingham Shum, S., Sándor, Á., Goldsmith, R., Bass, R., & McWilliams, M. (2017). Towards Reflective Writing Analytics: Rationale, Methodology, and Preliminary Results. *Journal of Learning Analytics*, 4(1), 58–84. <https://doi.org/10.18608/jla.2017.41.5>
- Christof, E., Köhler, J., Rosenberger, K., & Wyss, C. (2018). *Mündliche, schriftliche und theatrale Wege der Praxisreflexion: Beiträge zur Professionalisierung pädagogischen Handelns* (1. Auflage). hep der bildungsverlag.

- Combe, A. & Kolbe, F.-U. (2008). Lehrerprofessionalität: Wissen, Können, Handeln. In W. Helsper & J. Böhme (Hrsg.), *Handbuch der Schulforschung* (S. 857-875). Springer.
- Fund, F.-M. (2019). Unterrichtsnachbesprechungen in schulischen Praxisphasen. Eine empirische Untersuchung aus inhalts- und gesprächsanalytischer Perspektive. Springer.
- Hattie, J. & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research*, 77(1), 81–112. <https://doi.org/10.3102/003465430298487>
- Jantowski, A., Bartsch, A.-M., Limmer, J., & Gumz, E. (2010). *Evaluationsbericht zum Praxissemester im Jenaer Modell des Lehramtsstudiengangs – Belastungsempfinden und Studienzufriedenheit im modularisierten Lehramtsstudiengang unter Praxisbedingungen – Jena 2010*. Jena. <https://www4.uni-jena.de/unijenamedia/Downloads/einrichtungen/zfd/Bericht+PS+Ende+Kohorte+1+Endversion.pdf>
- Korthagen, F. A. & Kessels, J. (1999). Linking theory and practice: changing the pedagogy of teacher education. *Educational Research*, 28(4), 4–17.
- Lai, G. & Calandra, B. (2007). Using Online Scaffolds to Enhance Preservice Teachers' Reflective Journal Writing: A Qualitative Analysis. *International Journal of Technology in Teaching and Learning*, 3, 66–81.
- Mientus, L. (2023). *Reflexion und Reflexivität - Befunde reflexionsbezogener Dispositionen*. Universitätsverlag Potsdam. <https://doi.org/10.25932/publishup-61000>
- Mientus, L., Wulff, P., Nowak, A., & Borowski, A. (2023). Fast-and-frugal means to assess reflection-related reasoning processes in teacher training – Development and evaluation of a scalable machine learning-based metric. *Zeitschrift für Erziehungswissenschaften*. <https://doi.org/10.1007/s11618-023-01166-8>
- Mientus, L., Wulff, P., Nowak, A., & Borowski, A. (2021): ReFeed: computerunterstütztes Feedback zu Reflexionstexten - Ein Lehrkonzept zur Förderung der Reflexionskompetenz angehender Physiklehrkräfte an der Universität Potsdam. In: M. Kubsch, S. Sorge, J. Arnold, & N. Graulich (Hrsg.): *Lehrkräftebildung neu gedacht - Ein Praxishandbuch für die Lehre in den Naturwissenschaften und deren Didaktiken*, Waxmann. S. 160 - 165. <https://doi.org/10.31244/9783830993490>
- Nowak, A. (2023). Untersuchung der Qualität von Selbstreflexionstexten zum Physikunterricht. Entwicklung des Reflexionsmodells REIZ. In: M. Hopf & M. Ropohl (Hrsg.). *Studien zum Physik- und Chemielernen*. Logos.
- Nowak, A., Kempin, M., Kulgemeyer, C., & Borowski, A. (2019). Reflexion von Physikunterricht. In C. Maurer (Hrsg.), *Naturwissenschaftliche Bildung als Grundlage für berufliche und gesellschaftliche Teilhabe. Gesellschaft für Didaktik der Chemie und Physik, Jahrestagung in Kiel 2018* (Bd. 11, Nummer 1, S. 838–841). Universität Regensburg. [https://doi.org/10.1016/0742-051X\(94\)00012-U](https://doi.org/10.1016/0742-051X(94)00012-U)
- Poldner, E., van der Schaaf, M., Simons, P.R.-J., van Tartwijk, J., & Wijngaards, G. (2014). Assessing student teachers' reflective writing through quantitative content analysis. *European Journal of Teacher Education*, 37(3), 348–373.
- Römer, J., Rothland, M., & Straub, S. (2018). Bedingungsfaktoren des Beanspruchungs- erlebens von Lehramtsstudierenden im Praxissemester. In J. König, M. Rothland, & N. Schaper (Hrsg.), *Learning to Practice, Learning to Reflect?: Ergebnisse aus der Längs- schnittstudie LtP zur Nutzung und Wirkung des Praxissemesters in der Lehrerbildung* (S. 265–286). Wiesbaden: Springer VS.
- Roters, B. (2012). *Professionalisierung durch Reflexion in der Lehrerbildung*. Waxmann.
- Rothland, M. (2021). Die „Lehrerpersönlichkeit“: das Geheimnis des Lehrberufs?. *Die Deutsche Schule*, 113 (2021) 2, 188-198.
- Ullmann, T.D. (2019). Automated analysis of reflection in writing: validating machine learning approaches. *International Journal Artificial Intelligence in Education*, 29(2), 217–257.
- von Aufschnaiter, C., Fraij, A. & Kost, D. (2019). Reflexion und Reflexivität in der Lehrerbildung. Herausforderung Lehrer*innenbildung - Zeitschrift zur Konzeption, Gestaltung und Diskussion, Bd. 2 Nr. 1. *Herausforderung Lehrer_innenbildung*, (144- 159). <https://doi.org/10.4119/HLZ-2439>
- Wulff, P., Buschhüter, D., Westphal, A., Mientus, L., Nowak, A., & Borowski, A. (2022). Bridging the Gap Between Qualitative and Quantitative Assessment in Science Education Research with Machine Learning—A Case for Pretrained Language Models- Based Clustering. *Journal of Science Education and Technology*. <https://doi.org/10.1007/s10956-022-09969-w>
- Wyss, C. & Mahler, S. (2021). Mythos Reflexion. Theoretische und praxisbezogene Erkenntnisse in der Lehrer*innenbildung. *Journal für LehrerInnenbildung*, 21(1), 16- 25. <https://doi.org/10.35468/jlb-01-2021-01>