Author: Dr. Tim Fütterer

Hit the road towards an automated assessment of written reflections of pre-service teachers

Your full name:

Dr. Tim Fütterer

Affiliated authors with institutions:

Tim Fütterer[1], Kathleen Stürmer[1], and Richard Göllner[1]

[1] Hector Research Institute of Education Sciences and Psychology, University of Tübingen, Germany

Affiliation:

Hector Research Institute of Education Sciences and Psychology, University of Tübingen, Germany

Current position:

Postdoc

Title of your paper:

Hit the road towards an automated assessment of written reflections of pre-service teachers

Abstract

A promising approach to promote theory-practice integrated knowledge early in teacher education are written reflections of videotaped teaching situations. Pre-service teachers must apply their professional knowledge about aspects of teaching quality (e.g., goal clarification) to describe, explain, and predict observed classroom situations (reasoning; van Es & Sherin, 2002). Written reflections showing pre-service reasoning are indicators for their quality of professional knowledge (Seidel & Stürmer, 2014). Even though such reflective tasks are reasonble to connect theory and practice, no validated instrument exist to evaluate written reflections in an efficient way. In the project «An intelligent feedback system for observing instructional videos (INFER)», we developed a coding scheme to evaluate written reflections. Based on the coding scheme, we will use an artificial intelligence (AI) assisted evaluation to automate the evaluation and to provide immediate feedback to pre-service teachers on their quality of knowledge. To train the AI, the coding scheme includes prompts to stimulate varying levels of reflection quality, which in turn, will then be used to efficiently train the AI. We used prompts that differ in degree of guiding reflective processes (i.e., reasoning). As little is known about how written reflections should be segmented and coded by humans to train an AI reliably and validly, we conducted a study in which *N*=220 pre-service teachers reflect one to four staged videos of mathematics and language lessons (>800 written reflections). The videos are representative for the teaching quality aspects *goal clarification* and *feedback*. Currently, four intensively trained assistants systematically segment (sentence and sense units) the written reflections and code the segments regarding reasoning. We will present initial results on rater agreements regarding differences between segmentation units, prompts, and aspects of reasoning. We will discuss the findings in the context of teacher education and provide further information to achieve an automated feedback system.

References

Seidel, T., & Stürmer, K. (2014). Modeling and measuring the structure of professional vision in preservice teachers. *American Educational Research Journal*, *51*(4), 739–771. https://doi.org/10.3102/0002831214531321

van Es, E. A., & Sherin, M. G. (2002). Learning to notice: Scaffolding new teachers' interpretations of classroom interactions. *Journal of Technology and Teacher Education*, *10*(4), 571–596.

Author: Dr. Tim Fütterer
Hit the road towards an automated assessment of written reflections of pre-service teachers

Extended Summary

Introduction

Worldwide, the ability to reflect on one's own practice is understood as a key characteristic of professional teachers (Darling-Hammond, 2021). The encouragement of teachers to engage in meaningful reflections about their teaching practice or teaching related attitudes and beliefes is assumed to be helpful to integrate achieved theoretical background into actual teaching practices (Korthagen, 2014). According to assumptions about professional competence acquisition, the linkage of conceptual knowledge acquisition with practice representations is required already at early stages of professionalization (Boshuizen et al., 1995). A frequently used approach to gain insights into and to promote reflective processes of teachers are written reflections (e.g., journal writing: Bain et al., 2002; portfolios: Körkkö et al., 2016). A promising and well-studied approach are written reflections of videotaped teaching situations. Videotaped teaching situations offer the advantage of systematically reducing the complexity of teaching situations, making them more analytically accessible to pre-service teachers (Grossman et al., 2009) and to promote theory-practice integrated knowledge early in teacher education. Systematically exploiting the potential of written reflections of videotaped teaching situations in teacher education requires an ecologically valid procedure whereby written reflections are evaluated in a timely manner, even for large groups of students. However, even though such reflective tasks are reasonble to connect theory and practice, no validated instrument exist to evaluate written reflections in an efficient way. In the project INFER, we developed a coding scheme to evaluate written reflections which will be used to train an artificial intelligence (AI) for automated evaluations.

Theoretical Background

Reflective processes are often described in different levels. Starting from the description, more elaborate levels of reflection up to critical reflections are defined, which have been formulated in numerous level models (e.g., Hatton & Smith, 1995; Korthagen & Vasalos, 2005). In our study, we use the prominent approach to assess the quality of reflections derived from the well-known concept of reasoning, which is formulated within the theory of professional vision (e.g., Stürmer & Seidel, 2015). Reasoning encompasses the processes describing (e.g., teaching situations), explaining and predicting. Describing implies identifying and differentiating various teaching events based on professional knowledge without making judgments. Explaining means that different observable teaching events are related to theories of learning-effective teaching. Predicting means that on the basis of learning theories possible consequences of the different teaching events for the students are estimated. When pre-service teachers are asked to reflect on classroom situations, then pre-service teachers must apply their professional knowledge about aspects of teaching quality to describe, explain, and predict observed classroom situations (reasoning; van Es & Sherin, 2002). Written reflections showing pre-service reasoning are indicators for their quality of professional knowledge (Seidel & Stürmer, 2014). Even though, these quality dimensions of reflective processes have been established in research on teachers professional vision, their practical use in teacher education is limited for two reasons. First, validated instrument to assess the quality of reflective processes about teaching are scarce. Existing instruments refer to more general aspects of self-reflections in different occupational contexts, and therefore, do not provide an assessment of teachers reflection competencies tailored to teaching. Second, the coding of teacher reflections to provide feedback is time-consuming (typically written reflections need to be manually coded by humans). In general, modern AI procedures can ensure an automated scoring of written reflections, and thus be useful to reduce time and effort. Moreover, AI-based evaluation systems are able to access all available information of a written reflection (vgl. Luckin et al., 2016). This represents a decisive potential and difference to the standardized video-based measurement instruments developed so far, which work with closed response formats (e.g., Observer

Author: Dr. Tim Fütterer

Hit the road towards an automated assessment of written reflections of pre-service teachers

Research Tool; Seidel & Stürmer, 2014) and therefore bound to the preselection of videotaped teaching situations. However, the existing AI systems need a tremendous amount of training material to achieve an acceptable accuranceny and precision of scoring. Furthermore, it is not sufficiently known how an AI can be trained in the best possible way with respect to quality criteria (i.e., reliability, validity). For instance, written reflections are often segmented as sense units (i.e., thematically related text sections). Sense units are preferred to sentence units because "undergraduate students are inclined to use long compound sentences to express multiple complex messages or summations in a single entry in their reflective journals" (Shek et al., 2021). However, systematically designed empirical studies on the benefits and effects (e.g., rater agreement, differences of subsequent coding) of different segementation approaches are scarce. Such information are important for reliable and valid trainings of an AI for automated evaluation of written reflections. Regarding these challenges, we systematically address research questions, starting with questions like:

1. How reliably can segments and aspects of reasoning be determined in written reflections?
2. How do different prompts affect the reliability of determining pre-service teachers reasoning in written reflections?

Method

To address such research questions, we implemented a highly standardized research approach. In a first step, we conducted a study in which *N* = 220 pre-service teachers were asked to reflect one to four staged videos (duration of 2 to 4 minutes) of mathematics and language lessons. The videos were selected in elaborate validation processes so that they are representative of two central teaching quality indicators of classroom management *goal clarification* and *feedback* (Seidel & Stürmer, 2014). Furhtermore, the videos are well structured and therefore represent deliberate and targeted reductions of the complexity of teaching situations (Grossman et al., 2009). As students need to be encouraged to write purposefully and meaningfully (Picard & Imhof, 2010; Zeichner & Liu, 2010), pre-service teachers were supported by prompts. More than 800 written reflection available; each written reflection at least 400 words long. Currently, four intensively trained assistants segment (sentence and sense units) the written reflections and code the segments regarding reasoning. All written reflections will be segmented twice, and each segment will be coded twice. Furthermore, we assessed several personal variables (e.g., personality traits, test scores regarding reasoning [Stürmer & Seidel, 2017]) that will be used for both predicting reasoning of pre-service teachers in written reflections and training the AI.

References

Bain, J. D., Mills, C., Ballantyne, R., & Packer, J. (2002). Developing reflection on practice through journal writing: Impacts of variations in the focus and level of feedback. *Teachers and Teaching*, *8*(2), 171–196. https://doi.org/10.1080/1354600220127368

Boshuizen, H. P. A., Schmidt, H. G., Custers, E. J. F. M., & Van De Wiel, M. W. (1995). Knowledge development and restructuring in the domain of medicine: The role of theory and practice. *Learning and Instruction*, *5*(4), 269–289. https://doi.org/10.1016/0959-4752(95)00019-4

Darling-Hammond, L. (2021). Defining teaching quality around the world. *European Journal of Teacher Education*, *44*(3), 295–308. https://doi.org/10.1080/02619768.2021.1919080

Grossman, P., Compton, C., Igra, D., Ronfeldt, M., Shahan, E., & Williamson, P. W. (2009). Teaching practice: A cross-professional perspective. *TeachersCollegeRecord*, *11*(9), 2055–2100.

Author: Dr. Tim Fütterer

Hit the road towards an automated assessment of written reflections of pre-service teachers

Hatton, N., & Smith, D. (1995). Reflection in teacher education: Towards definition and implementation. *Teaching and Teacher Education*, *11*(1), 33–49. https://doi.org/10.1016/0742-051X(94)00012-U

Korthagen, F. A. J., & Vasalos, A. (2005). Levels in reflection: Core reflection as a means to enhance professional growth. *Teachers and Teaching*, *11*(1), 47–71. https://doi.org/10.1080/1354060042000337093

Korthagen, F. A. J. (2014). Promoting core reflection in teacher education: Deepening professional growth. In C. J. Craig & L. Orland-Barak (Eds.), *International Teacher Education: Promising Pedagogies (Part A)* (Vol. 22, pp. 73–89). Emerald Group Publishing Limited. http://www.emeraldinsight.com/doi/abs/10.1108/S1479-368720140000022007

Körkkö, M., Kyrö-Ämmälä, O., & Turunen, T. (2016). Professional development through reflection in teacher education. *Teaching and Teacher Education*, *55*, 198–206. https://doi.org/10.1016/j.tate.2016.01.014

Picard, C., & Imhof, M. (2010). Prompts zur Anleitung tiefenorientierten Schreibens in Lerntagebüchern und Portfolios. In M. Gläser-Zikuda (Ed.), *Lerntagebuch und Portfolio aus empirischer Sicht* (Vol. 27, pp. 59–80). Verlag Empirische Pädagogik.

Shek, M. M.-P., Leung, K.-C., & To, P. Y.-L. (2021). Using a video annotation tool to enhance student-teachers' reflective practices and communication competence in consultation practices through a collaborative learning community. *Education and Information Technologies*, *26*(4), 4329–4352. https://doi.org/10.1007/s10639-021-10480-9

Stürmer, K., & Seidel, T. (2015). Assessing professional vision in teacher candidates: Approaches to validating the observer extended research tool. *Zeitschrift Für Psychologie*, *223*(1), 54–63. https://doi.org/10.1027/2151-2604/a000200

Stürmer, K., & Seidel, T. (2017). A standardized approach for measuring teachers' professional vision: The Observer Research Tool. In E. O. Schack, M. H. Fisher, & J. A. Wilhelm (Eds.), *Teacher Noticing: Bridging and Broadening Perspectives, Contexts, and Frameworks* (pp. 359–380). Springer International Publishing. https://doi.org/10.1007/978-3-319-46753-5_21

Zeichner, K., & Liu, K. Y. (2010). A critical analysis of reflection as a goal for teacher education. In N. Lyons (Ed.), *Handbook of Reflection and Reflective Inquiry: Mapping a Way of Knowing for Professional Reflective Inquiry* (pp. 67–84). Springer. http://link.springer.com/chapter/10.1007/978-0-387-85744-2_4