Combining and integrating formative and summative assessment in mathematics teacher education

Nils Frederik Buchholtz¹ · Nadine Krosanke² · Anna B. Orschulik² · Katrin Vorhölter²

Accepted: 25 May 2018 / Published online: 29 May 2018
© The Author(s) 2018

Abstract

Contrary to the opinion that formative and summative assessment approaches are not compatible, this article presents a theoretically grounded way in which different forms of assessment can be combined and integrated in university mathematics teacher education. Two mixed-assessment approaches are demonstrated through the analysis of a case study involving a practice-based seminar accompanying a school internship. First, a formative eportfolio assessment was combined with a summative panel survey to assess the learning opportunities of mathematics pre-service teachers. Second, the formative eportfolio approach was integrated with a summative oral course examination to make statements about the learning processes and learning outcomes of the pre-service teachers. Our analyses conclude that combining and integrating the two forms of assessment present the possibility of evaluating different aspects of the pre-service teachers' perceptions of opportunities to learn. Benefits, validation aspects and limitations of the two approaches of combining and integrating assessment forms are discussed.

Keywords Formative assessment · Summative assessment · Mixed assessment · Mixed methods evaluation · Eportfolio · Teacher education · School practice · Opportunities to learn

1 Introduction

Pre-service teachers in Germany frequently bemoan the lack of relation between their university studies and their later careers in the field, claiming that the courses they attended did not adequately prepare them for their work with students (Heublein et al. 2010). Therefore, in recent years, school-practical studies in different formats (e.g., internships) play a larger role as learning opportunities in German teacher education. Innovative university courses designed to accompany these field experiences provide pre-service teachers with possibilities of uniting theoretical knowledge and teaching practice. These courses aim to reduce the discontinuity between theoretical knowledge acquired at university and professional experiences in the classroom (Arnold et al. 2014; Zeichner 2010). One aim of the University of Hamburg's project ProfaLe ("Professional teaching to promote

subject-based learning under changing social conditions"; Kaiser 2015), for example, is to improve the seminars that are intended to complement school internships for pre-service teachers. The newly developed seminar structure takes up practical experiences and focuses on the development of situation-specific aspects of teaching competence. In this paper, we focus on the seminar for the mathematics pre-service teachers.

The advantages and opportunities for practical-based learning approaches have been discussed in teacher education for years (Putnam and Borko 2000). However, we as teacher educators have to ask ourselves: How can we evaluate the profit that pre-service teachers gain in terms of theoretical and practical expertise in their internship and in the accompanying seminar? Moreover, what are the appropriate structures and tools to use in order to assess the development of pre-service teachers' professional competence in such practical-based learning approaches? It is important to monitor the development of professional competence and pre-service teachers' opportunities to learn when accompanying pre-service teachers and supporting them on their way to becoming professional teachers.

✉ Nils Frederik Buchholtz
n.f.buchholtz@ils.uio.no

¹ University of Oslo, PB 1099, Blindern, 0317 Oslo, Norway

² University of Hamburg, Hamburg, Germany

Teachers' professional competence is, according to Blömeke et al. (2015), a continuum that comprises dispositions such as cognitions and affect-motivations as well as situation-specific skills and performance. In this paper, we argue that corresponding to the different aspects of teaching competence, different forms of assessment are beneficial when it comes to assessing the outcomes of pre-service teachers' learning processes. In general, two different approaches to assessment can be distinguished: *Summative assessment* (SA) measures the achievement of previously defined standards, tasks or goals, encapsulating all collected evidence up to a given point to yield either comparative or numerical ratings (Taras 2005). *Formative assessment* (FA), on the other hand, promotes individual development by interpreting and providing feedback according to a diagnostic judgment that presents information about the candidates' continuative learning processes. Unfortunately, many see the distinct purposes of these two forms of assessment as incompatible (William 2010). But in our understanding of SA and FA, these two approaches can be *combined* and even be *integrated* to assess the development of professional teaching competence, because each assesses different aspects of competence (e.g., knowledge about study content and the prevalence of perceived opportunities to learn about different study content), FA focuses in on the contextual situatedness of the pre-service teachers' professional actions and the development of situation-specific skills.

Our study aimed to traverse the classical dichotomy between SA and FA in mathematics teacher education. In this paper, we demonstrate a combination of SA and FA and an integration of FA and SA in an evaluation of the outcomes of a practical-based seminar designed to accompany school-internships in mathematics teacher education.

2 Theoretical considerations on using mixed-assessment approaches

In the frequently cited review of research on the assessment of competencies in higher education by Blömeke et al. (2015), the authors model competence as a continuum that comprises dispositions, situation-specific skills, and performance. This model can also describe the professional competence of mathematics teachers.

Dispositions contain, for example, cognitive components like the professional knowledge teachers acquire through university-based learning opportunities (Shulman 1986). *Situation-specific skills* comprise three different aspects related to practice: "(a) *Perceiving* particular events in an instructional setting, (b) *Interpreting* the perceived activities in the classroom and (c) *Decision-making*, either as anticipating a response to students' activities or as proposing alternative

instructional strategies" (Kaiser et al. 2015, p. 374; see also; Schoenfeld 2011). In this *Perception–Interpretation–Decision-making (PID) model*, it is assumed that the availability of situation-specific skills significantly determines whether the transformation from disposition to performance succeeds (Blömeke and Kaiser 2016). Since this model implies a rather holistic view of competence, Blömeke et al. (2015, p. 8ff) also point to the challenges of an appropriate assessment: Pure cognitive-analytical approaches might lack validity because relevant parts of the measured constructs can be underrepresented, and pure performance-based assessments might neglect the contribution of dispositional resources. Assuming that "the whole is greater than its parts" (p. 9), Kaiser et al. (2017) recommend working with a broader range of combined and situated assessment formats that are able to cover processes mediating the transformation of dispositions into performance. We agree that accurately and reliably measuring teaching competence requires a mixture of SA and FA assessment. This assessment approach can be described methodologically as either a combination or an integration of the different assessment forms. In the following, we clarify the difference between combination and integration and concretize how each approach addresses unique aspects of the professional teaching competence of our pre-service teachers.

2.1 Combining SA and FA forms

When evaluating pre-service teacher education measures and the outcomes of learning, SA generally must play a role, since any educator assessment aims at least to certify that certain course planning and teaching skills have been acquired. During the internship and our accompanying seminar, pre-service teachers are provided with opportunities for cognitive learning as well as for the acquisition of situation-specific skills. In our case, then, a SA might therefore cover pre-service teachers' cognitive knowledge dispositions and the theoretical learning opportunities provided by our seminar.

On the other hand, FA is often brought into play as a pragmatic option for the support of pre-service teachers in their practice. Tillema (2010) encourages assessments aimed at improving field-related performance:

Supporting teachers as learners requires moving beyond providing mere knowledge of results to a type of functional (i.e., performance-related) feedback that can articulate developments in practice, anticipate professional learning needs, and monitor the learner's progress during a course of teacher action (p. 563).

Tillema therefore proposes to "increase the learner control over the content to be assessed as well as over the criteria by which the teaching performance will be

scrutinized” (p. 567). Authors such as Bell and Cowie (2001) or Carr and Claxton (2002) also write from a socio-cultural view, understanding learning as a situated activity; these authors point to the situatedness of what they call *learning dispositions*. Accordingly, they favour FA approaches such as observations, interviews and self-reports like those that appear in portfolios when it comes to the assessment of performance-related competence. Thus, FA “values learning dispositions and sees [pre-service teachers’] early development as consequential on relationships between the learner and the social and material learning environment” (p. 18). To concretize this in our case, a FA might yield information about a pre-service teacher’s practical learning opportunities during his or her internship—in particular, the individual’s acquisition of situation-specific skills. Further, the FA provides us as evaluators with information about our pre-service teachers’ personal experiences with practical teaching, which we then might use, either to provide feedback, or adaptively for future seminar planning.

However, this formative approach faces methodological challenges. Unlike SA forms, FA forms presuppose that we have to give up the idea of comparing the achievements of pre-service teachers in a standardized way, since the learning progress is different for each individual pre-service teacher. FA requires evaluators to scrutinize whether all intended educational goals or certain measures in pre-service teacher education have been achieved. Because the identification of evidence of performance in FA is difficult, so far, this has been a critical issue when FA is used in teacher education or professional development contexts (Delandshere and Arens 2003). Accordingly, how mathematics (pre-service) teacher education can benefit from FA is so far not described very well (e.g., Spanneberg 2009).

Against this background, the basic idea of using a *combination* of SA and FA is to extend the scope of the assessment, i.e., to both widen and deepen the assessment by gathering as much information as possible on *multiple components of the assessed phenomenon* or *different, but closely related phenomena*. Assessment results should ideally complement each other to increase the overall interpretability of assessment results. For our case, the combination of SA and FA holistically addresses learning opportunities to increase teaching competence in order to gain information about the development of teaching competence in the internship and the seminar. Additionally, in favour of an increased interpretability of assessment results, we seek a mutual complementation of the findings from a standardized SA with the findings from an individual FA (e.g., when identifying evidence for learning opportunities in FA) and seek ways that we might partly overcome the challenge posed by the inability to standardize FA in teacher education.

2.2 Integrating SA and FA forms

According to Wiliam and Black (1996), neither FA nor SA categorically exclude each other; rather, they are seen as the extremes of a common continuum, the core of which is the (interpretable) evidence of performance:

Any assessment must elicit evidence of performance, which is capable of being interpreted (however invalidly). Whether or not these interpretations and actions satisfy the conditions for formative functions, the fact that interpretable evidence has been generated means that the assessment can serve a summative function. Therefore, all assessments can be summative (i.e. have the potential to serve a summative function), but only some have the additional capability of serving formative functions (p. 544).

Wiliam sees the integration of SA and FA as a combination of the different *purposes* of the assessment forms on the widest possible basis for evidence (Wiliam 2000, p. 11). In particular, for the use of SA as FA, further feedback from the assessor is necessary—based on the interpretation of the SA results and oriented toward helping the evaluatee continue on from assessed to desired levels of performance (see e.g., Sadler 1989; Taras 2005; Hattie and Timperley 2007). On the other hand, the idea of using FA as SA is to aggregate the separate results of a set of assessments designed to serve formative purposes in order to get a comprehensive picture of overall achievements (Harlen and James 1997).

The basic idea of an *integration* of SA and FA is therefore *to assess one phenomenon with different assessment forms*, for example by using small units of SA in a formative way in the course of giving feedback, and integrating this FA into an overall SA afterwards. For our case, this means that we are able to focus at the same time on how the pre-service teachers develop professional teaching competence and what outcomes of the seminar and the internship they achieve. In particular, the FA during the internship is composed of predetermined tasks affording particular situation-specific skills (described in Sect. 3.3), and we as lecturers provide constructive feedback on those tasks, helping the pre-service teachers to increase their teaching competence. The overall SA of our seminar—an oral examination conducted afterwards—individually addresses the situation-specific skills that the pre-service teachers have acquired during their internship. The aim behind this procedure is to increase the autonomy of the pre-service teachers concerning their certification of teaching competence on the one hand, while on the other hand increasing the validity of each assessment of situation-specific skills through the mutual corroboration of the results of FA and SA.

This approach is of course also challenging, since we as lecturers both enable learning processes and certify the

outcomes of the seminar and the internship. Shavelson (2006) highlights this “double bond” of the assessors of educational processes:

Significant tensions are created when the same person, namely the teacher, is required to fulfill both formative and summative functions. Teachers at the interface of formative and summative assessment ... confront conflict daily as they gather information on student performance—to help students close the gap between what they know/can do and what they need to know/be able do on the one hand, and to evaluate students’ performance for the purpose of grading/certification on the other hand (p. 8).

This double bind can only be circumvented with regard to the integration of FA and SA by not using or interpreting the same evidence for both purposes of assessment (William 2000). Within the seminar, we therefore differentiated clearly between learning situations (internship, seminar sessions and observation tasks) and achievement situations (oral examination) and made this clear to the pre-service teachers as well.

3 SA and FA in the ProfaLe seminar and corresponding research questions

The aim of the ProfaLe seminar was to develop the professional teaching competencies of pre-service teachers and to strengthen school-practical aspects, in order to reduce the common feeling of discontinuity between theoretical knowledge acquired at university and professional experiences in the field.

3.1 Structure of the seminar

The seminar was implemented across two periods of pre-service teachers’ practical experiences during their master’s studies. The first period was a full semester during which they split into pairs and spent 1 day per week in a school and 2 h weekly in a university seminar covering different topics related to mathematical teaching. The second period, which was held over the semester break, involved the completion of a 5-week internship, during which they—in the same pairs and schools—spent every day in school. Throughout both periods, the pre-service teachers observed and accompanied experienced teachers and even had to teach small units by themselves, supervised by the teachers and once by the university lecturers. In the seminar, the pre-service teachers learned how to focus their attention on important pedagogical classroom events that promoted their ability to perceive important classroom situations, interpret them, and decide how they should be managed. Therefore, the

structure of nearly all sessions of the seminar was based on the PID model (see Sect. 2). Initial points for analyses were, for example, observations of the pre-service teachers or videotaped classroom situations. A specific emphasis was placed on the heterogeneity of students and how teachers dealt with it.

Summing up, different forms of opportunities to learn were provided: First, the content of different seminar lessons provided opportunities to learn in the form of knowledge acquisition. With the help of this knowledge, the pre-service teachers were able to develop a certain teaching competence disposition. Second, observation tasks that the pre-service teachers had to carry out within their once-weekly internship during the semester provided another form of opportunity to learn: in this case, to perceive and interpret teacher behaviour. By doing so, they were asked to reflect on necessary teacher skills in certain situations. Third, within the oral examination of the seminar, the pre-service teachers were prompted to reflect on their own situation-specific skills concerning a self-chosen topic, focusing especially on the development of these skills during the internship. Thus, the oral examination also reflected opportunities to learn. Whereas the observation tasks involved reflecting on the skills of others, in the oral examination, the evaluatees were asked to reflect on their own skills.

In summer 2016 these innovations were implemented in the seminar for the first time in a group of 30 pre-service teachers (7 males, 23 females). Every pair of students was accompanied by a practicing teacher at their schools once per week during the semester and every day during a practical phase in the semester break for 5 full weeks (Orschulik 2016).

3.2 Implementation of SA

A special opportunity to integrate SA into the seminar was an interdisciplinary evaluation panel survey by the project ProfaLe that focused on the pre-service teachers’ professional teaching knowledge as well as on their university and school-practical learning opportunities (Doll et al. 2018). The survey draws on established instruments from the interdisciplinary study TEDS-LT (Blömeke et al. 2013; Buchholtz et al. 2016), a German follow-up to the international TEDS-M study released in 2008 (Tatto et al. 2012), and thus allows a standardized assessment of pre-service teachers’ professional knowledge and respective learning opportunities. University and school-practical learning opportunities—on which we focus in this paper—were surveyed in the form of an inquiry into perceived study contents, and thus they reflect dispositional aspects of teaching competence. The pre-service teachers were asked to indicate whether they had studied various content in their previous studies. The surveyed university

learning opportunities refer to six different subject-specific topics of mathematics education with a particular focus on aspects from the ProfaLe measures (see Table 1). The assessment was summative because the pre-service teachers were asked to recall a variety of different learning opportunities in the past and up to the point of the survey. Further details on the survey instruments are given in the paper by Doll et al. (2018).

The panel survey was conducted online and was open to all pre-service teachers of mathematics at the University of Hamburg. In total, 187 pre-service mathematics teachers took part. Participation in the online survey was voluntary on the basis of an honorarium and was anonymous due to data privacy regulations. However, since the online survey also asked about respondents' participation in specific university courses such as the ProfaLe seminar, we did obtain data concerning the participants of our ProfaLe seminar from the survey. In total, 13 of the 30 pre-service teachers from the ProfaLe seminar participated in the online survey.

A further form of SA for assessing teaching competence in the seminar was the oral examination at the end of the semester. It was administered as a presentation-and-reflection examination designed to reveal the outcomes of pre-service teachers' learning processes both during their school internships and in the seminar. This was accomplished by presenting a discussion about a self-chosen practical teaching example and included selecting important experiences from their own teaching activities or observations and linking them with the mathematics educational theory taught in the seminar. This way the pre-service teachers could stress relevant aspects within a 15-min presentation. But the oral examination did not serve only summative purposes; it also integrated formative aspects because the pre-service teachers had to choose the topic of their oral examination by themselves, based on their written reflections from their eportfolio. Furthermore, the examination had to center on a specific situational problem encountered during the internship. Prior to the examination, the pre-service teachers had to hand out a mind map to the lecturers, meaning a visual presentation

(graphical or textual) mapping their reflections about the chosen topic.

3.3 The implementation of FA

The FA of the pre-service teachers' professional teaching competencies was implemented in the seminar in a variety of ways, primarily through an electronic portfolio (eportfolio) and corresponding feedback, as well as through the self-responsible design of the oral examination presented above. A crucial part of the eportfolios was a set of observation tasks about which the pre-service teachers had to write.

Generally, eportfolios can be used in different ways: The feedback on the written observations given by university teachers can be used by the pre-service teachers for their further learning processes as well as by their university teachers for designing appropriate further teaching (Hattie and Timperley 2007). The use of eportfolios in the field of teacher education in general can be of great benefit for professional development, as different case studies on the use of electronic blended learning in teacher education show (Mackey 2009; Gikandi et al. 2011; Vogel 2018). They are especially helpful when they are used in tandem with school practice; when they provide authentic practice-related assessment activities (e.g., realistic classroom situations); and when they receive effective and prompt formative feedback (Gikandi et al. 2011, p. 2338). Still, although the use of online FA tools like eportfolios has increased in the last decade (Jafari and Kaufman 2006), there are also methodological problems related to the possibility that portfolios may create inference for performance (Delandshere and Arens 2003).

We decided to use an eportfolio approach to focus on the development of situation-specific skills of teaching competence. The pre-service teachers received in total 18 situational-observation tasks, which they had to address during their observations of other teachers or their own teaching in school, and to work on as a written contribution to their personal eportfolio.

In the seminar, three different forms of observation tasks were used:

Table 1 Opportunities to learn in mathematics teacher education

Dimension	Content	Example	Items	Reliability (Rho)	Cronbach's α
Mathematics education	Basics of mathematics education	Didactics of algebra	20	0.86	0.84
	Dealing with heterogeneity	Language in mathematics education	8	0.62	0.62
	ICT in mathematics education	Learning with apps	6	0.66	0.56
	Methods of instruction	Problem-solving approaches	8	0.76	0.75
	Curriculum	National curricula	8	0.70	0.67
	Research on mathematics education	Role of beliefs in mathematics learning	9	0.72	0.68

1. Six tasks in advance of a forthcoming lecture, in which the pre-service teachers had to describe and partly interpret classroom situations. For example: “Describe when and how the teacher took up errors from the students in the classroom. What was done with these errors? What kind of problems occurred in these teaching situations?”
2. Seven subsequent tasks as a follow up of a lecture, in which the pre-service teachers not only had to describe but had to focus more on interpretation and developing alternatives to act. For example: “Choose one of the heterogeneity aspects that you want to look at more closely in your observations. Describe the students according to this aspect, and how this aspect has influenced measures of differentiation in the classroom. Evaluate what kinds of measures were more and less beneficial for the students.”
3. Five tasks of students’ own choice of observation focus, in preparation for the oral examination.

Formulating the observation tasks required the pre-service teachers to focus on one specific topic—e.g., dealing with errors—following the structure of the PID model. Thus, pre-service teachers had to describe the perceived situations, interpret them, and sometimes were prompted to develop useful responses to them.

All written contributions were read by the seminar lecturers over the following week and commented on within the personal eportfolio, which only the lecturers had access to. However, the written contributions were compulsory for the seminar and thus were extrinsically motivated reflections. (This duality depicts the difficult relationship between heteronomy and autonomy within this assessment format; See Sect. 2.2). As the reflections represent the learning processes of the pre-service teachers, no rating or grading took place. The seminar lecturers provided feedback constructively and showed possibilities for the pre-service teachers to improve their instructional quality in the described situations, according to the criteria already discussed in the seminar. By recognizing the benefit of the feedback for the pre-service teachers’ own professional development, we hoped for a high acceptance of the feedback (Topping 1998).

The reflections in the eportfolio aimed at self-assessment and raising awareness of the writers’ own situation-specific skills, so that explicit situation-specific learning opportunities and learning processes could be reflected upon in detail through the given feedback. In this way, the pre-service teachers were able to use the feedback from the seminar lecturers to optimize their observations in the subsequent weeks as well as to improve their own teaching (Vogel 2018). However, we as seminar lecturers also were able to benefit from this form of FA: We could not only examine the results of the pre-service teachers’ reflections on the teaching practices they observed and their own teaching but we could

also use adaptive planning to incorporate those observations into upcoming lectures. While the school internship took place over the semester break, the pre-service teachers had to choose observation tasks by themselves. In this way, the pre-service teachers could use their entries in the eportfolio as a preparation for their individual oral examination (which then served summative purposes). Thus, we were able to avoid the conflict of overregulating the pre-service teachers through fixed observation tasks, and instead allow them to develop their own sense of interpretation about their learning development—another step in this process during which we were able to witness the adaptivity of FA (Tillema 2010).

3.4 Research questions

To identify the benefits of the combination and integration of SA and FA in the case of the ProfaLe project of mathematics teacher education, we focus on the empirical results of the different assessment forms guided by two research questions:

1. What forms of evidence of pre-service teachers’ professional competence and learning opportunities can we identify from combining SA (panel survey) with FA (eportfolios)?
2. Which complementary forms of evidence can we identify when we integrate SA with FA in the oral examination?

To answer these questions, we first describe in the methodological considerations how we analyzed our data, within the scope of this paper. In the Sects. 5 and 6, we illustrate and discuss the results of the combination and the integration of the different forms of assessment.

4 Methods for data analysis

4.1 Methods of analyzing SA data

Of the 30 pre-service teachers in our seminar, 13 took part in the online survey, which allowed us to analyze their professional knowledge and their university- and field-based learning opportunities as a measure for dispositional teaching competence (cf. Sect. 3.2). To analyze the learning opportunities, a Mokken analysis (Mokken and Lewis 1982), based on the whole sample of the online survey ($N=187$) was used. This non-parametric item response model allows us to check the assumption that the learning opportunities of different dimensions can be described hierarchically along an ordinal scale, so that the percentage of perceived content can be interpreted as individual progress along a study pathway. In Table 1, the reliability-corrected rank correlations

(Spearman's Rho) of the six dimensions, as well as Cronbach's α , are reported. The scales showed a satisfactory reliability in the range of 0.62–0.86. Since an analysis on the individual level was not possible due to data regulations, for our analysis, the group of the 13 pre-service teachers from the ProfaLe Seminar was compared to a comparative group of 28 pre-service mathematics teachers stemming from the whole sample. This group of pre-service teachers were enrolled in mathematics education courses at the same University in the same phase of study, but they did not participate in any teacher education courses affected by the ProfaLe project. We used the non-parametric Mann–Whitney U test for the comparison of medians and interquartile ranges of the perceived learning opportunities.

The SA of the oral examination was carried out according to the following overarching criteria, which were communicated to all pre-service teachers beforehand, in order to provide the possibility of comparability of the performance:

- The observations and own experiences were selected reasonably and are meaningful;
- An appropriate selection of theories of mathematics education has been chosen;
- The observations and their own experiences are comprehensibly presented, and analyzed and evaluated with the aid of theory;
- Changes and learning processes are taken up;
- Current and relevant, in particular didactic and scientific literature is included;
- The proper use of scientific terminology is employed;
- The presentation is self-reliant, adequately prepared and structured;
- Questions from the lecturers can be answered;
- Links to different topics of the seminar can be drawn.

Each oral examination was rated by two university teachers according to the criteria listed above, directly following its conclusion. All examinations were logged.

4.2 Methods of analyzing FA data

We subjected all eportfolios to a systematic process of analysis following suggestions from Delandshere and Arens (2003) about identifying evidence of performance from portfolios. We therefore combined a deductive and an inductive approach introduced by Mayring (2015) in order to be as open as possible for the analysis. We coded extracts from the eportfolios according to the learning opportunities provided by the ProfaLe seminar and indicators of situation-specific skills in teaching competence that the pre-service teachers described in their observation tasks. The deductively created part of the coding system (see Table 2) is based on the work of Sherin and Van Es (2009). The category “Stance”

was modified so that the three situation-specific skills could be identified, whereby the situation-specific skill of perception could be coded by utterances at the descriptive level. We changed the codes belonging to the category “Topic” into those that were topics of the seminar. By doing so, we were able to analyze to what extent these topics provided opportunities to learn in the form of knowledge acquisition, and sometimes of situation-specific skills. An example for a coding process can be found in Fig. 1 in Sect. 5.

In addition, new indicators for learning opportunities emerged inductively from the eportfolios, according to which the extracts also could be rated, e.g., descriptions about the development of the personality as a teacher or the professional relationship between the pre-service teachers and their mentors. The collection and classification of these extracts showed the overall heterogeneity of the learning opportunities and situation-specific skills of the pre-service teachers. Therefore, in the following section, we showcase some illustrative examples from pre-service teachers who have given us permission to present their work.

5 Results

To answer our research questions, we first present results of the combination of SA and FA, then the results of the integration of SA and FA. Subsequently in the Sect. 6, we discuss the results on a more theoretical level and also reflect on the limitations of our approach.

5.1 Results of the combination of SA and FA

With the results from the online survey, we can arrive at the number of perceived university learning opportunities concerning specific study content, and thus create an indicator for professional competence. However, we are able to analyze the outcomes of learning opportunities only on the group level. The findings are therefore complemented and deepened by those from the FA on the individual level, which are also used to make conclusions about the application of the study contents with regard to the development of situation-specific skills. Table 3 describes the results of the Mann–Whitney U test for the comparison of medians and interquartile ranges (IQRs) of the online survey.

With regard to the perceived study contents, we find significant differences from the comparison group. Thus, the group of pre-service teachers from our seminar felt they were presented with a significantly higher number of learning opportunities in the areas of dealing with heterogeneity and research on mathematics education. For example, the median of 0.38 in dealing with heterogeneity means that, on average, pre-service teachers from our seminar stated that they had previously encountered learning opportunities for

Table 2 Deductive part of the coding system

Coding categories	Coding
Actor	Student Teacher Other
Topic	MC 1—mathematical argumentation MC 2—mathematical problem solving MC 3—mathematical modelling MC 4—using mathematical representations MC 5—operate with symbolic, formal and technical elements in mathematics MC 6—Communicating Dealing with errors Mathematical concepts (“Grundvorstellungen”) Changing and connecting representations Language awareness in learning mathematics Cognitive activation Problem orientation Teacher intervention Differentiation and individualized instruction Form of teaching—student-centered Form of teaching—teacher-centered Classroom management
Stance (Situation-specific skills)	Describe Interpret Decide

MC Mathematical competency

three of the eight subthemes on dealing with heterogeneity (e.g., dyscalculia, giftedness). Despite the lack of significant differences in other dimensions, we recognized the high perception of study content related to “methods of instruction” and the comparatively low perception of content of “Information and communication tools (ICT) in mathematics education,” which could be expected from the seminar. These findings reflect the measures taken by the ProfaLe project, since dealing with heterogeneity and instructional methods constituted a major focus in the seminar (whereas the use of ICT was not in focus). In the following section, we combine these SA results with complementary individual results from the FA. Therefore, we concentrate on how the seminar and the internship offered learning opportunities for the development of situation-specific skills of the pre-service teachers, especially concerning dealing with heterogeneity.

A preparatory eportfolio entry on the topic of dealing with errors from Jennifer (Fig. 1), a female pre-service teacher (see Sect. 3.3), provides an illustrative example of an individual FA result. It is about observing the behaviour of a teacher dealing with an occurring misconception.

When describing her perceptions concerning the subject matter of the observation task, Jennifer focuses on dealing with errors. It becomes very clear that Jennifer observed the teaching and the classroom discourse very closely, focusing on both students and teacher. Her selection regarding this situation involves a nice example of an occurring misconception of inverse proportionality. Although not required, she interprets her perceptions, taking up other topics that were subjects of previous discussions in the seminar. A closer analysis of the “Interpret” and “Decide” segments shows that Jennifer does not explain the occurring error. However, since this was a preliminary observation task, it might not yet have been possible for Jennifer to rely on broad knowledge for error analysis. Furthermore, no error interpretation was expected according to the formulation of the observation task. Nevertheless, we can recognize that she already considers the teacher’s handling of the mistake as inappropriate and mentions the importance of students’ understanding. This reveals a high degree of pedagogical reflection on the learning of concepts in mathematics education, indicating her situation-specific skills when addressing

Fig. 1 ePortfolio entry from Jennifer and respective coding (translated by the authors)

Table 3 Median differences in perceived study contents

Group	n	Basics of MEd	Dealing with heterogeneity	ICT in MEd	Methods of instruction	Curriculum	Research on MEd
ProfaLe seminar	13	0.55 (0.30)	0.38* (0.25)	0.33 (0.42)	0.75 (0.19)	0.50 (0.25)	0.33* (0.17)
Comparative group	28	0.58 (0.24)	0.25 (0.22)	0.33 (0.33)	0.75 (0.25)	0.43 (0.38)	0.22 (0.11)

* $p < 0.05$, IQRs in brackets

misconceptions, and that might reflect her theoretical knowledge in this field.

Following the idea of FA, these and observations from other pre-service teachers were used by the lecturers to adapt teaching and learning in the course. In the next seminar session, observations and interpretations of the

pre-service teachers were used to discuss occurring errors of students on the basis of concrete examples and to consider how teachers can deal with misconceptions in the classroom. Thereby our aim was to accumulate the practical expertise of the pre-service teachers adaptively—as

intended in FA—and to develop guidelines for handling mistakes in the classroom together.

In summary, the online survey provided us with general information about pre-service teachers' perceptions of their own opportunities to learn, while the reflections from the observation tasks provided a good learning opportunity for training in situation-specific skills. This was the case not only when the pre-service teachers had the opportunity to observe examples of good teaching practice but also when they observed examples that called for improvement.

With our coding approach in the scope of this paper, we were able to make conclusions about the skills developed so far. However, our coding system requires closer examination if we are to make statements about the quality and appropriateness of the interpretations and decision-making. Nevertheless, when looking at individual cases, it is possible to show how perceived study contents are linked to other study contents and are available not only at the knowledge level but are also applied in describing, interpreting and decision-making. In this way, the results of the SA can be deepened and partially explained.

5.2 Results of the integration of SA and FA

Regarding the overall performance in the seminar, 28 of the 30 pre-service teachers took the oral examination at the end of the semester. The average grade was 1.5 and the grades ranged from 1.0 to 2.7 [grades may differ from 1.0 (best) to 4.0 (worst)]. The topics of the oral examinations varied. There were examinations about the handling of errors in the classroom, about certain measures of differentiation (e.g., circuit learning), cooperative learning in mathematics education or specific aspects of teaching methods, such as classroom discourse.

To illustrate how FA affected the topics in the SA for individual pre-service teachers, the following describes Leonie, a female pre-service teacher. One of her eportfolio entries relates to the topic “measures of differentiation in mathematics education” (see Example, Sect. 3) and describes how the teacher Leonie observed dealt with the heterogeneity of the class.

In the class the children work with a very different speed. About 3 or 4 students work very fast and usually are also correct. There is a larger mass of students working ‘mediocre fast’ and some who are very slow. (All get the same tasks, and all have to work on them at the same time.) Mondays, however, there is a learning time, which forces the children to work independently on either the subjects German, English or Mathematics in order to complete and/or practice tasks/content from the lessons. The students can use this time however they choose. Apart from that, there are sometimes

‘asterisk tasks’ as well for the children that finished very quickly. These tasks do not have to be done by all children. On the whole, I find this works out quite well. As a result of the learning time, even the slower children have the chance to work things off. I would find it great, however, if content would be offered that cognitively demands more of the faster children and leads them in some places to get deeper into a topic, to making connections, etc. I have the feeling there is not so much space to explore mathematical creativity. Possibly open task formats could be introduced for children who already finished. (I have not yet considered two children with learning disabilities in the class. I know that it is hardly possible for a teacher to provide additional meaningfully selected learning material which is connected to the current teaching, but these two children are just employed to utilize the ‘leftovers.’ I do not see any substantive thread there; no fostering plans have been consulted. I doubt that the math teacher knows what is in it. I think these children should also try to follow the current teaching, at a different level and at a different pace.) (Leonie, translated by the authors).

Leonie describes the teacher's measures in a very detailed way, taking an appraising attitude. The entry shows her development of situation-specific skills. It becomes clear that she perceives the situation in her observation in a holistic manner. The questioning of the cognitive activation and openness of the tasks shows that her situation-specific interpretations are to some extent also based on professional knowledge. Furthermore, she questions the teacher's actions against the background of the increased need for differentiation, especially in dealing with children with special needs. Although not required, she also develops alternative suggestions for differentiation measures, which are characterized by ideas of open- and closed-differentiation formats as well as by collaborative learning on a common object, that take the high-performing students as well as the two students with learning disabilities into account. Overall, her descriptions of the observations indicate a high degree of situation-specific skills including indications that she is already able to make informed decisions.

During the semester, the observation tasks in the eportfolios were not graded; rather, they served as learning opportunities. Accordingly, we gave the pre-service teachers learning feedback following the idea of FA. At the end of the semester, however, the oral examination is a graded certification exam that summatively assesses the pre-service teachers' professional teaching competence. Since the topic of the oral examination could be determined independently by the pre-service teachers and based on their own experiences, the oral examination also took up aspects of FA. The

integration of SA and FA therefore coalesces in the summative assessment of formative performance components (which, however, now are evaluated and graded in the examination according to the criteria mentioned in Sect. 4.1). We are able to integrate the findings from this kind of SA and FA on an individual basis. Identical situation-specific aspects of the pre-service teachers' professional competence can be found in their eportfolio entries, and in the topics for the oral examination on which the pre-service teachers elaborated in the mind maps, indicating that the examination assessed teaching competence that the pre-service teachers learned during the seminar and their internship.

One example of how the FA influenced the SA in that way appeared in Leonie's oral examination. In the course of her experiences during the internship, she continued to deepen the topic of measures for differentiation and individualization and chose this as the topic of her oral examination. The mind map she submitted clearly illustrates the questions and problems she addressed (see Fig. 2).

By integrating the results of the analysis of her eportfolio entries (recall Sect. 5.2) with her mind map, we can see that she did not stop at developing a critical attitude according to

her observations. In preparation for the oral examination, she deepened her knowledge and was able to relate her experiences to the—however implicit—theoretical background of the topic by dealing with legal requirements and finding further possibilities for differentiation measures stemming from research on mathematics education such as, for example, tasks for natural differentiation (Scherer and Krauthausen 2010). She also describes experiences from her own lessons, in which she used several differentiation measures to deal with the heterogeneity of the learning group—a demonstration of her situation-specific skills. It is also important to note that she critically reflects the limitations of differentiation measures based on her own and observed teaching practice.

6 Conclusions

6.1 Combining and integrating FA and SA

By combining different forms of assessment in the framework of a mixed-assessment approach, we hoped to

Theses:

1. In lessons, there should be phases in which students with learning difficulties work on the same topic as all the other students.
2. When students with learning difficulties work at the same topic as the other students, the type and the levels of the differentiation should be well-considered and a certain support should be available.
3. Differentiation has limits that are not least determined by the resources available.

Fig. 2 Leonie's mind map (translated by the authors)

determine the best possible evidence of pre-service teachers' learning opportunities and situation-specific skills. With SA (online survey) data, we were able to determine comparatively that the pre-service teachers of the seminar had significantly higher perceptions of the number of learning opportunities regarding heterogeneity and research in mathematics education that were available to them than were held by comparable pre-service teachers who did not take the seminar. However, these results alone are not a sufficient measure for the dispositional aspects of the pre-service teachers' professional competence since it is not clear whether collective differences could be attributed to enrolment in our seminar. We therefore used the data from the FA, which links the individual learning opportunities of the pre-service teachers with specific situational aspects of the seminar and the school internship. In this combination, the FA makes it possible to make detailed statements about the types of learning opportunities the pre-service teachers had and about how their learning processes differed. Both assessment forms provide us with indicators of the specific aspects of teaching competence of the seminar's pre-service teachers, thus providing a more comprehensive picture of their professional development.

While the *combination* of SA and FA focuses more on the mutual complementarity of assessment outcomes on different aspects of a phenomenon (in our case dispositional and situational learning opportunities), the *integration* of SA and FA is more likely to provide convergent findings about a phenomenon from different perspectives. In the eportfolios, we could identify pre-service teachers' situation-specific skills; these skills also formed the basis of their performance in the oral examination, so that the SA validly reflected the actual learning processes the pre-service teachers had. But the integration of SA and FA in the oral examination also provided a comparable framework for the assessment and certification of teaching competence for all pre-service teachers, regardless of the learning processes that took place during the internship. The situation-specific skills assessed in the FA could be certified by using SA criteria specified by the examination. By this means, a central challenge of FA in teacher education—namely the inability to standardize specific marks of evidence of performance—can be addressed. Furthermore, the integrated approach allowed not only situation-specific skills to be assessed but also knowledge components, so that professional teaching competence could be assessed with more validity. Thus, the two assessment approaches developed here—combination and integration—address different facets of the validity of assessments, which we discuss further below.

6.2 Addressing validity with mixed-assessment approaches

To understand the importance of distinguishing between a combination and an integration of SA and FA in terms of the validity of the approach, it is necessary to take our assessment to a broader theoretical level and analyze approaches to creating validity in qualitative or mixed-methods evaluation approaches (see also Frechtling and Sharp 1997; Gikandi et al. 2011). Since different methods entail different weaknesses and strengths, Denzin opted for a “methodological triangulation”, which consists of a “complex process of playing each method off against the other so as to maximize the validity of field efforts” (Denzin 1978, p. 304). But as in assessment, the attempt to triangulate different measures can lead to validity problems, because “different methods ... can relate to different empirical phenomena and ... it thus may be difficult to simply compare research results acquired by means of different methods in order to check their validity” (Kelle and Buchholtz 2015, p. 332).

This pragmatic view—which is also leading the debate on mixed-methods evaluations, e.g., in health science (Johnson and Schoonenboom 2016)—can stimulate an alternative understanding of triangulation that can also be transferred to the context of assessment in teacher education or education in general. The use of different forms of assessment to elicit and interpret evidence of performance can be compared with the examination of a physical object from two different viewpoints or angles. Both viewpoints provide different pictures of this object, which may or may not be useful to validate each other, but regardless may yield a fuller and more complete picture of the object if brought together. So, if SA and FA are mixed in any way, the following validity-related outcomes can arise (see also Kelle and Buchholtz 2015, p. 333):

- The elicited items of evidence of performance converge,
- The elicited items of evidence relate to different aspects of the performance, but are complementary to each other and thus can be used to supplement each other,
- The elicited items that are evidence of performance are divergent or contradictory, or
- The elicited items of evidence refer to unrelated phenomena.

For multiple reasons, then, it makes sense to combine or to integrate SA and FA within a mixed-assessment approach, whether to increase or enhance the validity of the evaluation (Johnson and Schoonenboom 2016). An integration of SA and FA may lead to convergent evidence and thus to valid interpretations, or divergent evidence and validity problems. We encountered these patterns in our analysis of the outcomes of a seminar in teacher education when we integrated the pre-service teachers' reflections on their field

observations from the FA with their SA-oriented oral examinations. In particular, we analyzed how their reflections influenced the determination and the preparation of the topic of their oral examination. We identified convergent evidence, which indicated that the SA really did assess teaching competence that was gained during the seminar and the internship, and thus validly assessed teaching competence. When assessing a combination of SA and FA, different aspects of the performance may yield complementary evidence of performance, or it may yield unrelated evidence. In our case, an online survey about pre-service teachers' perceived learning opportunities revealed an overview about perceived study contents on the group level. We could deepen these aggregated findings and partially explain them with individual findings from the FA despite the fact that the online survey and the eportfolios focused on different kinds of learning opportunities. The analysis of the eportfolios revealed complementary evidence with regard to perceived learning opportunities and how the study contents were related and applied in pre-service teachers' teaching practices.

Whatever way SA and FA are mixed, the main rationale behind the use of different forms of assessment is always the attempt to compensate for limitations of one form of assessment by drawing on the strengths of another form (see also Kelle and Buchholtz 2015). The decisive factor here is whether and how SA and FA can be related to each other (or can be mixed) and what type of evidence of performance the respective use of an assessment form produces. The form of the mixing needs to be described and justified based on the individual case.

6.3 Limitations of our approach

There are certain limitations in our approach that we have to take into consideration when reflecting on the benefits and challenges of a combination or an integration of SA and FA. In the combination approach, the indicators of learning opportunities are different because of the different assessment forms. An example of this aspect is how the topic *heterogeneity* influenced the development of teaching competence. While the results of the SA provided aggregated but standardized data (e.g., about the perceived learning opportunities with regard to heterogeneity), the analysis of the FA data is conducted on an individual basis and is still subject to subjective interpretations despite our methodological approach (Delandshere and Arens 2003). Even if the topic *heterogeneity* was a substantive part of the observation tasks, we were not able to reconstruct in every eportfolio entry how this topic affected the development of pre-service teachers' professional competence. The usefulness of a mixed-assessment approach can in general only be estimated based on individual cases and with regard to the assessment purposes. However, this limitation applies

to any form of assessment. Furthermore, our combination of FA and SA does not allow any mutual validation of the results of the assessments. Integration of the results of the learning opportunities to strengthen the validity of results from SA (online survey) and FA (eportfolios)—as is usually intended in mixed-method approaches—would have been accomplished only if the results of the FA could have been interpreted directly against the background of the results of the SA. For this to happen, participation in the panel survey would have to be made obligatory and non-anonymous, but we refrained from that due to ethical and data-protection reasons.

Acknowledgements We thank Prof. Dr. Jörg Doll for providing the data from the ProfaLe online survey.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Arnold, K. H., Gröschner, A., & Hascher, T. (Eds.). (2014). *Pedagogical field experiences in teacher education: Theoretical foundations, programmes, processes, and effects*. Münster: Waxmann.
- Bell, B., & Cowie, B. (2001). *Formative assessment and science education*. Dordrecht: Kluwer.
- Blömeke, S., Bremerich-Vos, A., Kaiser, G., Nold, G., Haudeck, H., Keßler, J.-U., & Schwippert, K. (Eds.). (2013). *Professionelle Kompetenzen im Studienverlauf: Weitere Ergebnisse zur Deutsch-, Englisch- und Mathematiklehrerbildung aus TEDS-LT*. Münster: Waxmann.
- Blömeke, S., Gustafsson, J. E., & Shavelson, R. (2015). Beyond dichotomies: Competence viewed as a continuum. *Zeitschrift für Psychologie*, 223(1), 3–13.
- Blömeke, S., & Kaiser, G. (2016). Understanding the development of teachers' professional competencies as personally, situationally and socially determined. In D. J. Clandinin & J. Husu (Eds.), *International handbook of research on teacher education* (Chap. 52). Thousand Oaks: Sage.
- Buchholtz, N., Scheiner, T., Döhrmann, M., Suhl, U., Kaiser, G., & Blömeke, S. (2016). *TEDS-shortM. Kurzfassung der mathematischen und mathematikdidaktischen Testinstrumente aus TEDS-M, TEDS-LT und TEDS-Telekom*. Hamburg: University of Hamburg.
- Carr, M., & Claxton, G. (2002). Tracking the development of learning dispositions. *Assessment in Education: Principles, Policy and Practice*, 9(1), 9–37.
- Delandshere, G., & Arens, S. A. (2003). Examining the quality of the evidence in pre-service teacher portfolios. *Journal of Teacher Education*, 54(1), 57–73.
- Denzin, N. K. (1978). *The research act: A theoretical introduction to sociological methods*. New York: McGraw-Hill.
- Doll, J., Buchholtz, N., Kaiser, G., König, J., & Bremerich-Vos, A. (2018). *Nutzungsverläufe für fachdidaktische Studieninhalte der Fächer Deutsch, Englisch und Mathematik im Lehramtsstudium*.

- Die Bedeutung der Lehrämter und der Zusammenhang mit Lehrinnovationen. *Zeitschrift für Pädagogik*. (in press).
- Frechling, J., & Sharp, L. (Eds.). (1997). *User-friendly handbook for mixed method evaluations*. Arlington: National Science Foundation, Division of Research, Evaluation, and Communication.
- Gikandi, J. W., Morrow, D., & Davis, N. E. (2011). Online formative assessment in higher education: A review of the literature. *Computer and Education*, 57(1), 2333–2351.
- Harlen, W., & James, M. (1997). Assessment and learning: Differences and relationships between formative and summative assessment. *Assessment in Education: Principles, Policy and Practice*, 4(3), 365–379.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112.
- Heublein, U., Hutzsch, C., Schreiber, J., Sommer, D., & Besuch, G. (2010). *Ursachen des Studienabbruchs in Bachelor- und in herkömmlichen Studiengängen*. Hannover: HIS. (Forum Hochschule F2(2010)).
- Jafari, A., & Kaufman, C. (Eds.). (2006). *Handbook of research on ePortfolios*. New York: Idea Group Reference.
- Johnson, R. B., & Schoonenboom, J. (2016). Adding qualitative and mixed methods research to health intervention studies: Interacting with differences. *Qualitative Health Research*, 26(5), 587–602.
- Kaiser, G. (2015). Professionelles Lehrerhandeln zur Förderung fachlichen Lernens unter sich verändernden gesellschaftlichen Bedingungen (ProfaLe). BMBF-Qualitätsoffensive Lehrerbildung in Hamburg. In Presentation at the University of Hamburg. <https://www.profale.uni-hamburg.de/forschung/vortraege/qib---uk-jahre-stagung-vortrag-29062016.pdf>. Accessed 29 May 2018.
- Kaiser, G., Blömeke, S., König, J., Busse, A., Döhrmann, M., & Hoth, J. (2017). Professional competencies of (prospective) mathematics teachers: Cognitive versus situated approaches. *Educational Studies in Mathematics*, 94(2), 161–182.
- Kaiser, G., Busse, A., Hoth, J., König, J., & Blömeke, S. (2015). About the complexities of video-based assessments: Theoretical and methodological approaches to overcoming shortcomings of research on teachers' competence. *International Journal of Science and Mathematics Education*, 13, 369–387.
- Kelle, U., & Buchholtz, N. (2015). The combination of qualitative and quantitative research methods in mathematics education: A "mixed methods" study on the development of the professional knowledge of teachers. In A. Bikner-Ahsbals, C. Knipping & N. Presmeg (Eds.), *Approaches to qualitative research in mathematics education. Examples of methodology and methods* (pp. 321–361). Dordrecht: Springer.
- Mackey, J. (2009). Virtual learning and real communities: Online professional development for teachers. In E. Stacey & P. Gerbic (Eds.), *Effective blended learning practices: Evidence-based perspectives in ICT-facilitated education* (pp. 163–181). Hershey: ICI Global.
- Mayring, P. (2015). *Qualitative Inhaltsanalyse. Grundlagen und Techniken*. Weinheim: Beltz.
- Mokken, R. J., & Lewis, C. (1982). A nonparametric approach to the analysis of dichotomous item responses. *Applied Psychological Measurement*, 6, 417–430.
- Orschulik, A. (2016). Strengthening cooperation within school practical activities as a way to improve teacher education. In Poster presented at the 13th International Congress on Mathematics Education (ICME-13) in Hamburg. <https://www.profale.uni-hamburg.de/forschung/postervortraege/poster-orschulik-2016-icme.pdf>. Accessed 15 June 2017.
- Putnam, R. T., & Borko, H. (2000). What do new views of knowledge and thinking have to say about research on teacher learning? *Educational Researcher*, 29(1), 4–15.
- Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18, 145–165.
- Scherer, P., & Krauthausen, G. (2010). Natural differentiation in mathematics—The NaDiMa project. *Panama Post*, 3, 14–16.
- Schoenfeld, A. H. (2011). *How we think: A theory of goal-oriented decision making and its educational applications*. New York: Routledge.
- Shavelson, R. J. (2006). On the integration of formative assessment in teaching and learning with implications for teacher education. In Paper prepared for the Stanford Education Assessment Laboratory and the University of Hawaii Curriculum Research and Development Group. https://web.stanford.edu/dept/SUSE/SEAL/Reports_Papers/On%20the%20Integration%20of%20Formative%20Assessment_Teacher%20Ed_Final.doc. Accessed 29 May 2018.
- Sherin, M. G., & Van Es, E. A. (2009). Effects of video club participation on teachers' professional vision. *Journal of Teacher Education*, 60, 20–37.
- Shulman, L. S. (1986). Those who understand. Knowledge growth in teaching. *Education Researcher*, 15, 4–14.
- Spanneberg, R. (2009). The mathematics teaching portfolio: A reflective tool for developing professional growth and improving classroom practices. In B. Clarke, B. Grevholm & R. Millman (Eds.), *Tasks in primary mathematics teacher education. Purpose, use and exemplars* (pp. 51–65). New York: Springer.
- Taras, M. (2005). Assessment—summative and formative—some theoretical reflections. *British Journal of Educational Studies*, 53(4), 466–478.
- Tatto, M. T., Schwillie, J., Senk, S., Ingvarson, L., Peck, R., & Rowley, G. (2012). *Teacher education and development study in mathematics (TEDS-M): Policy, practice, and readiness to teach primary and secondary mathematics. Conceptual framework*. East Lansing: Teacher Education and Development International Study Center, College of Education, Michigan State University.
- Tillema, H. (2010). Formative assessment in teacher education and teacher professional development. In B. McGaw, E. Baker & P. Peterson (Eds.), *International encyclopedia of education* (pp. 563–571). Oxford: Reed Elsevier.
- Topping, K. (1998). Peer assessment between students in colleges and universities. *Review of Educational Research*, 66, 249–276.
- Vogel, R. (2018). Portfolioarbeit in der Grundschullehrerbildung—“Lernraum” zur Anbahnung eines professionellen mathematikdidaktischen und mathematischen Unterrichtshandelns. In R. Möller & R. Vogel (Eds.), *Innovative Konzepte für die Grundschullehrerbildung im Fach Mathematik* (pp. 197–218). Wiesbaden: Springer.
- Wiliam, D. (2000). Integrating formative and summative functions of assessment. In Paper presented to Working Group 10 of the 9th International Congress on Mathematics Education (ICME 9) in Tokyo. http://www.dylanwiliam.org/Dylan_Wiliams_website/Papers_files/ICME9%20WGA10%20paper.doc. Accessed 29 May 2018.
- Wiliam, D., & Black, P. (1996). Meanings and consequences: A basis for distinguishing formative and summative functions of assessment? *British Educational Research Journal*, 22(5), 537–548.
- Zeichner, K. M. (2010). Rethinking the connection between campus courses and field experiences in college- and university-based teacher education. *Journal of Teacher Education*, 61(1–2), 89–99.