Assessing the Outcome of Teacher Education Programs in Norway: An Analysis and Discussion of the Factor Structure in Domains of Teacher Practicum for Student Teachers at Three Norwegian Universities

Trond Solhaug, Thomas Dahl
Program for Teacher Education (PLU), Norwegian University of Science and Technology (NTNU), Trondheim, Norway
Email: trond.solhaug@plu.ntnu.no, thomas.dahl@plu.ntnu.no

Received 27 May 2016; published 16 July 2016; published 19 July 2016

Copyright © 2016 by authors and Scientific Research Publishing Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY).
http://creativecommons.org/licenses/by/4.0/

Abstract

This article contributes to our understanding of teacher education by presenting the results of a large pilot study of quality assessments of teacher education programs at three Norwegian universities. We present an assessment instrument that corresponds to the California Standards for Teaching Professions and the practice domains developed by Darling-Hammond. The research question is: To what extent does the reported five-factor structure in the reported revised instrument (Darling-Hammond) fit the Norwegian context? We have revised this instrument to better cover both the documented dimensions of teacher professionalism and the demands regarding teacher quality in Norwegian educational policy. This has been accomplished by incorporating and modifying the items of the Darling-Hammond instrument. The modification process involved a selection of students, teachers, teacher educators and a school headmaster who discussed the relevance and content validity of the items. The final questionnaire was piloted at three Norwegian universities in three different education programs, with the participation of 419 students. The results show high inter-correlation between the original practice domains of Darling-Hammond, but the original factor structure is reproduced and to some extent supported. The results are discussed, and we conclude that the reported assessment instrument is valid as a measure of the teacher education program contributions to students’ competence in the five reported domains of the teacher practicum. Therefore, the instrument may be used as an overall evaluation of the complete program or may be slightly modified using the item format to assess distinct aspects of teacher education.

Keywords
Teacher Education, Assessment, Confirmatory Factor Analysis, Norway

1. Introduction and Research Aim

This article addresses the development of an instrument to assess how well teacher education programs at three Norwegian universities prepare student teachers for their work as teachers. There are several reasons for this study.

First, there is a lack of overall quantitative instruments and assessments in teacher education in this country.

Second, student teachers are often repeatedly confronted with a challenge in school during their first practice for which they do not find themselves prepared. Sometimes, this is referred to a “reality shock” (Korthagen, Loughran, & Russell, 2006). Such a dramatic first teaching experience reflects a large gap between teacher education and its attempts to prepare students for the challenges they meet in today’s school (Brouwer & Korthagen, 2005; European Commission, 2013; Stokking, Leenders, De Jong, & Van Tartwijk, 2003) and that teacher education programs in many places in Europe are subject to change (Donaldson, 2014).

Third, in the aftermath of the Programme for International Student Assessment (PISA), the quality of teachers and teacher education has been debated and targeted for change, particularly in countries that perceive their PISA scores as unsatisfactory (Hökkä & Eteläpelto, 2014) (Darling-Hammond, Newton, & Wei, 2010; Korthagen, 2010).

Fourth, a major challenge in teacher education has been the problem of integrating academic subjects, particularly subject didactics (why, what and how to teach a subject in school), and more general pedagogic content knowledge (Finne et al., 2011; Korthagen, 2001; Lid, 2013; NOKUT, 2006) Accompanying this challenge is the ongoing debate on the irrelevance of some aspects of teacher education and its ability to meet students’ needs (Korthagen, 2006). Korthagen questions the logic of the transfer of theoretical knowledge from campus teaching to “student teachers’ practice in schools” (Korthagen, 2001).

Fifth, in evaluating the “New Public Management” (NPM), the trend has shifted focus toward the evaluation of outcomes. In particular, Cochran-Smith points out an historical shift toward the evaluation of outcomes at the turn of the millennium (Cochran-Smith, 2001), which is also the case for Norway (Solhaug, 2011).

Having pointed out these issues, the present article discusses an instrument to assess teacher education’s contribution to students’ performance as novice teachers. The focus of the instrument is on the perceived contribution of teacher education programs to novice teachers’ performance of various professional tasks as teachers.

There is no Norwegian standard for the teacher profession, but according to a government document, the curriculum for teacher education should meet seven different domains of competence (Kunnskapsdepartementet, 2009; Kunnskapsdepartementet, 2010). These are: 1) subject knowledge and five basic skills, 2) reflections on the relationship between school and society, 3) professional ethics as teachers, 4) preparation for good teaching in pedagogy and subject matter didactics, 5) classroom management and support for learning processes, 6) good cooperation and communication skills and 7) coping with change and professional development. In the present article, we regard these highlighted points as a framework for teacher education in Norway. We consider this to be an overall national competence framework for teacher education.

Choice of Assessment Instrument and Research Question

Because teacher education covers multiple sources of knowledge and practices, Cochran-Smith has argued for “multiple measures for assessing program outcomes” (Cochran-Smith, 2001; see also Darling-Hammond, 2006a). According to Darling-Hammond, one should evaluate the learning that occurs through particular courses as well as the program as a whole, the teaching performance of the student teachers as preservice candidates and novices and the outcomes of their performance (Darling-Hammond, 2006a: p. 123). Building on this, we decided to start our approach to the evaluation of the overall contribution of our program to student teachers’ performances as teachers. In so doing, we decided to build on Darling-Hammond’s analysis of student teachers in the Stanford Teacher Education Program (STEP) at Stanford University in California, USA. We argue that the
five dimensions of teacher professionalism reported as “factors” in this instrument (1, design of curriculum and instruction; 2, support for diverse learners; 3, use of assessment to guide learning; 4, create productive classroom environments and 5, teacher professional development) are also well in line with the Norwegian governmental policy framework. These dimensions correspond closely with the different divisions in the California Standards for the Teaching Professions (CSTP) (Darling-Hammond, 2006a).

In our choice of instrument, we emphasise that the STEP instrument is published, peer-reviewed and has proven to be scientifically valid. Further arguments for our choice are as follows. First, the instrument and item format focuses on the programs’ contribution to students’ actions and practical performance as teachers, which fits our assessment purpose. Second, the instrument covers professional aspects, such as planning, supporting learning, teacher assessments and teachers’ work vis-à-vis the learning environment in the classroom, all of which are at the heart of teaching. Third, it emphasises equal opportunities for all students, which implies that adapted learning is particularly important. Equal opportunities for all learners are also strongly emphasised in Norway (Kunnskapsdepartementet, 2007). Fourth, the diversity of pupils and the multicultural aspects of teaching are stressed, as these aspects are increasingly challenging in our (Norwegian) schools. Fifth, assessments for learning are greatly emphasised, which is also the case in Norway (Engh, Dobson, & Høihilder, 2007). Sixth, the classroom learning environment is strongly emphasised, which is also the case in Norwegian policy and professional practice. Seventh, critical reflection on teaching, teacher self-reflection and professional development are well in line with students’ right to participate.

Our research question is: To what extent does the reported five-factor structure in the STEP instrument fit the revised assessment instrument in the Norwegian context? (STEP, see Darling-Hammond, 2006)

We respond to this question by analysing the factor structure and an alternative factor structure reported in a revised instrument presented to student teachers in Norwegian universities. Empirically, the article is based on data gathered from student teachers in different teacher education programs at three different universities in Norway, with the participation of 419 students. We analyse data using the statistical software IBM-AMOS and perform confirmatory factor analysis. In order to discuss the model fit (Kline, 2005) versus the construct validity (Cook & Campbell, 1979; Kline, 2005), of our instrument, we also attempt a single factor solution and use the AMOS software’s modification index to feed the discussion statistically.

There are three types of programs at the three universities that are framed by 60 ECTS (The European credit transfer and accumulation system) and divided into three equal parts—pedagogy, subject matter didactics taught on campus and supervised practice in school. Although the three main programs may vary both within and between the universities with respect to curricular content, teaching and supervision in school, the main difference between them is about how the teaching programs are organized. The traditional program is taught as a full-time year-long additional study program after completing all exams from the bachelor or master programs. In the new “lector programs”, students sign up for the teacher education program from day one. The 60 ECTS of the three components are then taught as courses and/or periods of teaching practice in four out of five years. All students complete their masters in this program. These are the main programs at the universities. The third type of program is part-time to serve working teachers who need to qualify for teaching (schools may employ people on condition that they qualify within three years). These programs are usually taught on a part-time basis in a two-year period. The overall aim of the development of our assessment instrument is to be able to assess all these programs and compare their outcomes.

2. Previous Research

In the assessment of teacher education, there has been a focus on student teachers’ knowledge base (Cochran-Smith & Zeichner, 2010). The discussion of which facets this knowledge base contains and should contain has frequently been addressed in the literature. Blömeke et al. summarise this literature by stating that it is generally agreed to divide the knowledge into three different facets, two that are subject-related and one that is generic: content knowledge (CK), pedagogical content knowledge (PCK; including curricular knowledge) and general pedagogical knowledge (GPK) (Blömeke, Buchholz, Suhl, & Kaiser, 2014).

However, there is overwhelming scientific literature on teacher professionalism that clearly shows that teachers’ work includes far more domains than only subject and pedagogical knowledge. The most important and most addressed domain is professionalism and professional development. Several concepts have been developed to address this issue, such as professional learning communities (DuFour, 2004), professional capital (Andy
Hargreaves & Fullan, 2012) and collective teacher efficacy (Goddard, Hoy, & Hoy, 2000). Recently, John Hattie expanded his meta-analysis of what affects student achievement and found “collective teacher efficacy” to have the second strongest effect of a total of 195 influences (Hattie, 2015). The importance of professional interaction is also acknowledged at the policy level. When the European Commission searches for ways to improve the quality of teaching, they do not criticise teacher education for lacking quality in terms of knowledge building but its ability to support teacher professionalism. The European Commission states that one of the most important goals of teacher education is “to inspire teachers to be proactive, reflective professionals who take ownership of their own professional development” (European Commission, 2013: p. 35).

The idea of the school as a “learning organisation” has a strong foothold in Norwegian policy documents (Utdannings-og forskningsdepartementet, 2004), pointing to the need for teachers to cooperate and be involved in change and learning processes (Postholm et al., 2013). Organisational learning is also targeted as an important competence for school development and the development of education in research (Collinson, Tanya, & Sharon, 2006; A. Hargreaves, 2000; Scribner, Cockrell, Cockrell, & Valentine, 1999; Silins, Mulford, & Zarins, 2002; Simpson & Marshall, 2010).

In the Norwegian context, several investigations have been carried out to determine whether the schools are operating as learning organisations and if teacher education prepares them to work on school development. One study has validated an instrument to measure “organisational learning” (Postholm et al., 2013). Another study has validated an instrument to measure “professional identity”, “relevance for work as a teacher” and “development competence” in teacher education programs (Finne et al., 2011). Both these studies have items that explore the dimension of “developing professionally”, such as items dealing with doing research on one’s own practice, participating in development work and being able to reflect critically on policy issues. All in all, the Norwegian context and research call for the development of assessment instruments and a modification of the items laying the ground for the five different dimensions in Darling-Hammond’s questionnaire, as well as additional items.

3. Theoretical Framework for Teacher Education Programs

We acknowledge that there are scientific and educational differences between American traditions and practices and European traditions and practices. We therefore deliberately chose to present two frameworks for teacher education programs. Our approach in this section is to begin with two summaries of knowledge based on studies of seven American and three Dutch (European) teacher education programs. From the scholarly summaries of these programs we establish a framework for our discussion of factor model adjustments. The first is a summary of research provided by Linda Darling-Hammond and colleagues, and the second is provided by Dutch scholar Fred Korthagen and colleagues. These frameworks reflect empirical research about seven American and three Dutch teacher education programs. Darling-Hammond formulates a summary of characteristics of good education programs, while Korthagen summarises a study of teacher education programs on “principles guiding teacher education programs”. The frameworks are not directly comparable, but they reflect the views of teacher education programs and approaches to successful teacher education programs or the current adjustment and research of teacher education program outcomes.

The first framework deals with the “how” of teacher education and was developed by Darling-Hammond based on a study of seven successful teacher education programs. Learning-to-teach programs necessitate (according to Darling-Hammond, 2006b) addressing the fact that teaching may be quite different from student teachers’ own experiences. Student teachers need to be prepared not only to think but also to act as teachers, and student teachers need to be prepared to handle the ever-changing complexity of classrooms. She writes:

1st a common, clear vision of good teaching that permeates all course work and clinical experiences, creating a coherent set of learning experiences; 2nd well-defined standards of professional practice and performance that are used to guide and evaluate course work and clinical work; 3rd a strong core curriculum taught in the context of practice and grounded in knowledge of child and adolescent development and learning, an understanding of social and cultural contexts, curriculum, assessment, and subject matter pedagogy; 4th extended clinical experiences…supervised practicum…that are carefully chosen to support the ideas presented in simultaneous, closely interwoven course work; 5th extensive use of case methods, teacher research, performance assessments, and portfolio evaluation that apply learning to real problems of practice; 6th ex-
plicit strategies to help students to confront their own deep-seated beliefs and assumptions about learning and students and to learn about the experiences of people different from themselves; 7th strong relationships, common knowledge, and shared beliefs among school-and university-based’ faculty jointly engaged in transforming teaching, schooling, and teacher education (Darling-Hammond, 2006b: p. 305/6).

Darling-Hammond strongly emphasises consistency (relative consensus) between actors in various parts of theory and practice support in teacher education programs. There are strong beliefs in terms of competencies and assessments and shared beliefs and consensus regarding program design and curriculum.

The Dutch scholar Fred Korthagen (2006) seems more cautious in listing student teacher competences. In discussing the essence of a good teacher, he ultimately avoids responding to the question of what makes a good teacher by providing a list of competences and skills. In his approach to teacher education programs, he asks the research question: “What central principles shape teacher education programs and practices in ways that are responsive to the expectations, needs and practices of teacher educators and student teachers?” (Korthagen et al., 2006). Korthagen et al.’s (2006) list of principles is derived from the research and development of three programs combined and a significant amount of other research. His first principle is: “Learning about teaching involves continuously conflicting and competing demands” (Korthagen, 2006: p. 1025). The principle addresses the ability to handle complexities; conflicting demands need to be not only theorised about but also to be experienced and developed in practice. Conflicting demands are less explicit in the Darling-Hammond summary above but may be reflected in her sixth point. Korthagen’s second principle is: “Learning about teaching requires a view of knowledge as a subject to be created rather than as a created subject.” This principle puts much emphasis on student teacher practice theory as a summary of reflective experience. Korthagen advances a reflective model of learning from experience, action, looking back on action, awareness of essential aspects, create alternative methods of action and trial (ALACT). Darling-Hammond seems to put more emphasis on knowledge as a created subject as opposed to knowledge constructed on the basis of student teacher practices. It is at the very least unclear what the status of “created knowledge” (previous theory and research) is in relation to the knowledge developed through reflective practices in Darling-Hammond’s framework. Korthagen’s third principle is: “Learning about teaching requires a shift in focus from the curriculum to the learner.” This implies that learning to teach is embedded in the practices of students, which is an argument for situated learning in teaching. There is also a balance of power between curriculum as a given framework and the construct of teaching practice situated in a learning environment. Korthagen’s fourth principle is in line with the principle of reflectivity and the focus on the learner: “Learning about teaching is enhanced through (student) teacher research” (Korthagen, 2006: p. 1030). A distinction between reflectivity and research might be that research implies systematic collection and analysis followed by some form of documentation. In teacher education, this activity underpins the knowledge development of one’s own practice and construction of knowledge and theory (ising). The principle of research is less explicit in Darling-Hammond’s work. Research is also strongly emphasised in Finnish teacher education; for example, see Malinen, Väsänen, and Savolainen (Malinen, Väisänen, & Savolainen, 2012) and Tirri and Ubani (Tirri & Ubani, 2013) Korthagen’s fifth principle is: “Learning about teaching requires an emphasis on those learning to teach working closely with their peers” (ibid, p. 1032). Collaboration is not only logical from a constructivist and a sociocultural knowledge viewpoint but also from an organisational perspective in school. Sharing ideas and having an open school climate is always valued as constructive in knowledge institutions like schools. Above all, shared collegial experiences are often vital in order to improve classroom practice. The need for collaboration and open discussions is also highlighted in Darling-Hammond’s research (sixth point). Korthagen’s sixth principle is: “Learning about teaching requires meaningful relationships between schools, universities and student teachers.” This principle addresses the challenge in teacher education that actual practice and educational institutions need to be integrated to be informed and provide relevant teaching to students. Such collaboration is also emphasised in Darling-Hammond’s seventh point above. Learning about teaching is enhanced when the teaching and learning approaches advocated in the program are modelled by the teacher educator in their own practice. As pointed out by (Bandura, 1997), modelling is the best way to convey patterns of behaviour and attitudes to various aspects of (in this case) the professionalism of teaching. This aspect is less valued in the Darling-Hammond summary.

Despite differences in the frameworks, they both attempt to outline important aspects of teacher education programs.
Summary of Teacher Professionalism

There is no doubt that both scholars emphasise the need for knowledge as a foundation for professional teaching practice. Darling-Hammond emphasises consistency and relative agreement on the knowledgebase. Korthagen points out that knowledge is a necessity for good teaching but realises that transferring that knowledge from the educational institution is a problem. Therefore, every student teacher needs to construct their own knowledge in practice.

In the globalised world and the information age, there is increasing human diversity in school classrooms and in knowledge. Teaching practices and decisions should be grounded in solid knowledge gained from both education and good practice, which is reflective. Regardless of the diversity of teacher education, we assume that knowledge of, learning processes, motivation, assessments, child development, self-determination, Bildung, student diversity, human relations, ethics and curricular framework are at the heart of teachers’ knowledge base in their profession. These aspects of knowledge contribute to the performance of different tasks in the planning and design of teaching; supporting learning processes for diverse learners in class; managing assessment favourably, in particular support for learning processes; creating a constructive and supportive learning environment and supporting professional development.

Darling-Hammond has a strong focus on institutional consistency as an agreed-upon framework for teacher education, while Korthagen strongly emphasises teacher capability (and autonomy) in dealing with complexity of all kinds (and we agree). We also agree with him that collegial/parental collaboration and teacher research in their professional development should be part of teacher education programs and professional teaching. However, collaboration and professional development need to be grounded in solid professional knowledge (above). We recognize that both scholars address the need for institutional integration or meaningful relationships between various institutional aspects of teacher education.

4. Methodology

4.1. Sampling and Pilot Data

Three universities in Norway contributed to the pilot study on the condition that they (except for NTNU) were anonymous in any publications. Data were collected via questionnaire sheets distributed and collected by teacher educators and scanned at NTNU. A sample is given in Table 1 below.

At NTNU, students come from three different teacher education programs—a five-year program (127 students), a one-year full-time program (94 students) and a two-year part-time program (45 students). At the two other universities, students come from the traditional one-year program. The three programs are organised differently in terms of teaching and practice and may differ between universities. There may also be differences in curricula, teaching and practice, but all programs are equivalent to 60 ECTS. A total of 419 students (261 females, 62% and 158 males, 38%) at the three different universities responded to our questionnaire, which is a good sample for a pilot study. Because this is a pilot study, the sample size and variety of programs and universities is fairly good.

4.2. Analytical Departure

The theoretical framework above and the five dimensions (statistical factors) Darling-Hammond revealed through an exploratory factor analysis on data from student teachers at Berkeley is the basis for the current empirical research on factor structure (see also the research question in section 1). However, we have modified and added a number of items to adapt to local Norwegian language and local teacher education practice. In addition to the outlined theory, we are also indebted to Einar and Sissel Skaalvik for their work in revising and adding

<table>
<thead>
<tr>
<th>Table 1. Pilot data sample.</th>
</tr>
</thead>
<tbody>
<tr>
<td>University</td>
</tr>
<tr>
<td>University 1</td>
</tr>
<tr>
<td>University 2</td>
</tr>
<tr>
<td>NTNU</td>
</tr>
</tbody>
</table>
items to the STEP instrument. Their six-factor “Norwegian Teacher Self-Efficacy Scale” covers the area of teacher professionalism, which is highly reflected in the above theoretical framework. At the same time, their items pose interesting teaching challenges which extend the perceived teaching difficulties in the instrument. We have used a few of these items (Skaalvik & Skaalvik, 2007: p. 624).

4.3. Modifying Procedures and Content with Respect to the Norwegian Context and Research

An established working group suggested using a revised version of the STEP questionnaire, and the decision was made by the current leaders. A working group of three lead by the authors invited all colleagues in current teacher education programs to make suggestions regarding the revising of the first translated questionnaire. In addition, headmasters and some teachers representing various fields of practice were asked to do the same. The working group collected responses and discussed, selected and revised the items. A total of six versions of the questionnaire were developed before a revised questionnaire (paper) for electronic scanning was developed with the help of local senior engineer. Another 14 rounds of corrections/changes resulted in the present version. In the revision process, we particularly paid attention to verbs that reflect actions and to tasks in practice domains (Gable and Wolf 1993). In the following, we comment on the major changes, leaving out minor changes that aim at adapting item text to local Norwegian language use (see Appendix).

Section 1, “To develop and plan lessons”, is more or less kept as is, with only minor adaptations to fit local Norwegian language. Section 2 is “To lead and support different students’ learning”. Here, three aspects of teacher professionalism and students learning were added, which are all reflected in the theoretical framework above. First, we made explicit that teachers are in fact the leaders of learning processes in class, and we added item 2_V1. Second, we added the teacher responsibility to engage in motivating/encouraging students and included item 2_V2. Third, we added teacher responsibility not only to support individual learners but also to promote students’ social development both in their design of teaching and in their engagement with students (Items 2_V7 and 2_V9).

Section 3 is “To use assessment to guide and support for learning”. We added the social aspect of assessment in item 3_V3. In Norway, student participation is strongly emphasised in the law (Lov om grunnskolen og den vidaregåande opplæringa, 1998) (Education Act). Based on this, we added two items that reflect student participation and interactions with teachers in teaching and learning (Items 3_V6 and V7). Section 4 is “To create a good learning environment”. We replaced items which first of all address “trust” as a vital quality of relations in school. Norway is a trusting society, and schools depend on having a trusting culture. We added item 4_V1. Support for students and their good results are most important in teacher profession, which is reflected in Items 4_V2 and V3. The learning environment is always affected by challenging students, which is reflected in Items 4_V4 and V5. Here, we are indebted to Skaalvik and Skaalvik (2007). Section 5 is “Working with teacher professional development.” The original four items in this section are limited with respect to what research on teacher professionalism has revealed. Organization of Economic Cooperation and Development (OECD) has made a fruitful distinction between “teaching competences” and “teacher competences”, pointing to practices and self-development (OECD, 2009). In a recent document for the European Commission, “teacher competences” is defined as being based on “a wider, systemic view of teacher professionalism, on multiple levels—the individual, the school, the local community, professional networks” (European Commission, 2013, p. 10). With reference to Korthagen above, we added two items on teachers doing their own research (5_V2 and V9). Furthermore, we added the social aspect of professional development in Items 5_V4, V5, V10 and V12. We particularly emphasise the perspective of learning organisations here. We also added the ability to reflect critically on professional, ethical and policy issues, which is well in line with the professional aspects outlined above and which is emphasised in the local teacher education programs. At the end of the process, the questionnaire consisted of 46 items. The number of items in Darling-Hammond’s study was 29.

4.4. Data analytical Procedures

We started by analysing the distributions using kurtosis and skewness indexes. All distributions were well below the threshold of 2 (Christophersen, 2012). To check for context sensitivity of the item format and content, separate analyses revealed systematic and sometimes significant differences between the universities. This is a clear sign of sensitivity to different educational contexts. Missing values were from 0.2% to 1.5%, which is very low and is due to the careful data collection procedure. These missing values were replaced with mean of the indi-
individuals’ scores on the particular item. We have required that the individual should respond to at least 50% of the questions within each aspect. Ten individuals were removed in total, which is very low and a sign of high-quality data.

We performed a confirmatory five-factor analysis using the IBM SPSS AMOS program. In addition, alternative one-factor models were performed and are presented in this article. The first model was then modified according to the principles below.

Modifying a model to obtain better fit between the model and the empirical covariance matrix requires specific strategies. In our approach to the first model, we departed from the theory of teacher professionalism and empirical research. Our first principle is that the theory of teacher professionalism and related concepts has priority over empirical indications of model improvement. The IBM SPSS AMOS program (along with other such programs) provides such a modification index that suggests how particularly residual inter-correlations might improve the model fit. We used this index in our improvements, but we only accepted improvements for which we found good theoretical rationales or which do not challenge the conceptual validity of the instrument. This way, our approach does not challenge our guiding principle of priority for theory. Modification indexes often suggest specifying correlations between residuals. There are two ways to improve the model based on these suggestions. The first is to specify the residual inter-correlation in the model, and the second is to remove one item that is a discussion of simplicity of model versus validity claims. Residuals are unique variances (not part of the factor), and inter-correlations between these are a principal threat to the factor structure. We therefore decided to give priority to a simpler model and will follow this guiding principle as long as it does not seriously challenge concept validity. The modification index provided a large number above its threshold of 4 (Kline, 2005). This caused a major dilemma regarding whether to keep items in the model to secure the best possible content validity or to take items out or specify inter-correlations to get better model fit vis-à-vis the empirical matrix. We give priority to content validity in our analysis. A decisive argument for this is that items represent different aspects of teacher practice. The overall practical aim of this instrument is to assess how teacher education programs contribute to the different aspects of teacher professionalism and practice. The more aspects represented, the more are we able to get indications of improvements. A third principle is to start modifying the model one step at the time, always starting with the largest value in the modification index. In addition, we have particularly been looking for correlations between aspects of teacher education professionalism or program contributions that are a “threat” to the present factor structure.

5. Results

Figure 1 shows our first model.

5.1. Comments on Results

To begin, we provide a Cronbach’s $\alpha$ (Crocker & Algina, 1986) for each of the five factors: $V_1 \alpha = .88; V_2 \alpha = .90; V_3 \alpha = .85; V_4 \alpha = .86$ and $V_5 \alpha = .87$. The $\alpha$ scores are fairly good and show consistency. However, the $\alpha$ is also a product of the number of items, and some low factor loadings are moderately reflected in the scores. The overall fit of the model is not very good, is below the threshold of the RMSEA of .08 for acceptable fit and is below the threshold for NFI, CFI and GFI for good fit (Kline, 2005). There are strong inter-correlations between the five latent variables ranging from .80 to .88. Such correlations indicate that the five aspects of program contributions have much in common and possibly that the strong inter-correlation between the latent factors is one reason for the lack of model fit. Looking at the factor loadings, all items display reasonable loadings above .50, which is fairly good for so many items. The inter-correlation and moderate fit has encouraged us to assess an alternative model with only one factor in Figure 2 below.

This one-factor model has a worse fit than the five-factor model with a higher RMSEA of .084 and low values for all three indexes—NFI, CFI and GFI. Our preliminary conclusion from this is that the five-factor model still fits better, and we continue to modify this in the following. Recall our strategies for modification above. However, we will also carefully consider whether the lack of fit in the one-factor model is caused by inter-correlations between the latent factor residuals. One specific additional aim in the modification process is therefore to see whether the removal of such inter-correlations in the five-factor model may give better fit in the one-factor model.
Figure 1. Program contributions. First five-factor model of teacher education program quality assessment. Latent variables: v1, preparation and planning; v2, lead and support learning; v3, assessment for learning; v4, create a good learning environment and v5, professional development. Items refer to the questionnaire in the appendix. Numbers are correlations between latent variables, factor loadings and residual variances.
Figure 2. Program contributions. First one-factor model of teacher education program quality assessment. Numbers are factor loadings and residual variances.
5.2. Revised Models

We made six modifications to the five-factor model to obtain a better fit (see appendix for item references). First, there are inter-correlations between residuals in 4.V4 and 4.V5. We decided to keep 4.V5 because this item covers a broader substantial perspective. Second, 5.V3 and 5.V4 have a residual correlation. We decided to remove 5.V3 as the most specific item. Third, there is a residual correlation between 2.V4 and 4.V7. We kept 2.V4 because it covers the relevant professional aspect better than 4.V7. The latter item is also misplaced in the professional development construct. Fourth, 4.V3 and 5.V7 residuals correlate. We decided to remove 5.V7 because it is unclear about the school-related conflicts and because the correlations are across the factor structure in the model. Fifth, regarding the correlation between 3.V1 and 3.V2, we decided to keep 3.V2 because 3.V1 displays several inter-correlations and therefore does not seem to fit in very well. Sixth, research is covered in 5.V2, and 5.V9 covers a broader field of personal development. The revised models are presented in Figure 3 and Figure 4 below.

The improved model shows that we are able to obtain acceptable/good fit in terms of the RMSEA of .059, but the fit is still clearly below the threshold of .90 for the NFI and GFI indexes and but reaches almost .90 for the CFI. It is, however, quite noteworthy that the inter-correlation between the latent factors of professional domains is as strong as in our first model.

The threshold for significant improvements in the modification index is Chi^2 below 3.8. The empirical index provides quite a number above 4 which is the threshold value for reported possible corrections in the index. There are still a number of improvements to be made according to the index, but to do these one has to severely compromise content validity and/or increase model complexity. We have therefore decided to propose a conceptually valid model and will discuss the factor structure and the fit after presenting the revised one-factor model. Recall that we would like to explore whether the removal of items and inter-correlations between residuals across factors in the five-factor model would result in better fit. We therefore apply the same modification of items in model 4 below as in model 3 above.

The revised one-factor model has worse fit that the revised five-factor model for all indexes. This supports (moderately) the five-factor structure in model 3. To some extent, it also supports the original structure presented by Darling-Hammond (Darling-Hammond, 2006b). We discuss this issue below.

6. Discussion

The first issue is simply to ask whether it is meaningful to consider these broad practice domains as distinct statistical factors. To ask such a question is to ask whether the practice domains may be regarded as theoretical counterparts to the empirical items. On the one hand, we argue that our empirical tests lend some support to the five-factor model in the Norwegian context and to the five dimensions revealed by Darling-Hammond. A five-factor model clearly fits better than the one-factor model and supports the view that the practice domains reported here may be treated as distinct aspects of practice that are also reproduced in empirical tests. Theoretically and practically, planning, leading learning activities, assessments, working on learning environments and professional development are distinct domains of practice and reflection. Our suggested model should therefore work well in assessing and developing these domains in teacher education.

On the other hand, we argue that the strong inter-correlations between the latent five aspects in the five-factor model, regardless of modification attempts, support the fact that teaching is complex and that practice domains are fairly integrated. To illustrate this very briefly, in planning and development there are always experiences from classes and learning, assessments and different social environments. In carrying out plans in teaching and interacting in classes, there is always the issue of adjusting plans and drawing upon reflections (in action) of personal experiences and development. Professional development requires reflections on all aspects of one’s own practice. There are more examples, but suffice it to say that the professional domains of practice are also fairly integrated. This integration is strongly supported by the correlations between the latent aspects of professional teacher practice and some inter-correlation between residuals. We acknowledge that a good fit (RMSEA < .05, Chi^2 indexes > .90) might be obtained with much fewer items or other items but not at the expense of content validity or of the measurement of the practice domains. We therefore argue that the acceptable fit of the model is partly due to the broad theoretical and blurred distinction between the domains. It is partly due to our priority of conceptual validity and simplicity (avoiding specifying residual correlations) in the model. The factor loadings are fairly even and above .50. Note that a factor loading of .50 implies its commonality, h^2 = .25. In
Figure 3. Program contributions. Second modified five-factor model of teacher education program quality assessment. Latent variables: v1, preparation and planning; v2, lead and support learning; v3, assessment for learning; v4, create good learning environment and v5, professional development. Items refer to the questionnaire in the appendix. Numbers are correlations between latent variables, factor loadings and residual variances.
Figure 4. Program contributions. Second one-factor model of teacher education program quality assessment, revised. Numbers are correlations between latent variables, factor loadings and residual variances.
other words, 25% of the items’ total variance contributes to the factor, which is still a substantial contribution. In comparison, Darling-Hammond (2006) reports lower factor loadings (in the 40s) in her study, and a model fit is not reported. A preliminary conclusion is therefore that the presented five-factor model is meaningful and clearly better than the one-factor model.

The item format invites students to assess the contribution of all three parts of the complex teacher education program to various aspects of their practice. Recall there is an emphasis on program contribution to perform actions as teachers in the item format. To assess this connection between an educational program and one’s own performance is a complex task for a student, and this complexity might be a threat to the reliability of data. In support of the format, we argue that there are very good distributions in the data well below critical thresholds of skewness and kurtosis. We also observed in the data from our own university that students rated research on their own practice very favourably with a rather small standard deviation (below 1.0 on a 7-point scale measurement), which is also strongly emphasised in our program. Some other aspects of our program were much less favourably rated, which came as no surprise to us. Interestingly, these distributions and their often large standard deviations reveal large differences in teacher students’ perceptions of their teacher education (Solhaug, Dahl, & Holter, 2014). In addition, there are apparent differences between the universities in the data, which we assume reflects that the instrument is responsive to institutional variation.

In the modification process, we noted substantial logic in some of the residual inter-correlations. Also, surprisingly, few of the really large modification index numbers were due to residual inter-correlation between professional aspects. Based on the above data and research observations, we argue that this item format is quite responsive and appropriate for assessment use. A final comment on the item format is that when developing items, several theoretical and practical aspects must be carefully considered. First is the issue of which verbs (or actions) should be listed in the items. Second, there are everyday fields of practice and “acts of teaching” that should be represented; the issue here is the careful choice of which situations should be listed. Third, we discovered from the analysis of our pilot data from NTNU that the variety of challenges students face in their teaching practice is a somewhat neglected aspect in the present instrument. In our version for 2015, we included more challenges inspired by there writing of a few of Skaalvik and Skaalvik’s teacher self-efficacy items (Skaalvik & Skaalvik, 2007). In this study, we use the instrument as an overall assessment of the program. This may be one overall legitimate target, but such use limits our conclusions regarding the program as a whole, which makes it more difficult to use the results as a basis for improving the three parts. We therefore recommend that similar but more specific questionnaires be used to assess specific parts. Such assessments allow for analysis and improvement at more specific levels of teaching and/or practice.

Another issue we wish to raise is the use of assessments of practice domains instead of, for instance, a more clear conceptual focus, such as program contribution to teacher efficacy, work on pupil motivation, developing pupils’ learning, promoting pupils’ self-regulation, promoting pupils’ mastery and teacher-efficacy beliefs and enhancing pupils’ overall self-esteem to help them reach their aspirations. We argue that the presented questionnaire is necessary and important in its focus on key domains of practice but is not sufficient as a basis for a comprehensive assessment of the program. We may develop another questionnaire based on the above-mentioned aspects and other concepts. Such an assessment would allow for a more precise conclusion and consequently for more specific improvements. We therefore see a professional domain-specific assessment and more conceptual-based assessment as complementary.

7. Conclusions

From the above theoretical framework, empirical analysis and discussion, we conclude that the reported instrument for assessing quality in teacher education programs in the specific teacher practice domains is proved to have good content validity in the fields it measured. We also claim that it has statistical validity in terms of acceptable fit and good factor loadings. We argue that some lack of fit is due to moderate but unspecified residual covariance. The covariance reflects the broad and integrated theoretical concepts of practice domains where a strict factor structure may not be expected.

The instrument measures five different dimensions that contribute to the quality of teacher education. We set out to use dimensions that correspond with those explored by Darling-Hammond (Darling-Hammond, 2006a) and that mirror the California Standards for Teaching Professions. We have introduced new items and have modified some existing items in a number of fields. Referring to our methodological section, our new items
contributed to the following important aspects of teacher work and professionalism: teachers as leaders of learning processes; teacher as motivators; teaching as social development and social learning; teachers as facilitators of student participation; students perceived trust in relations; teaching in challenging situations; research as a professional requirement; critical policy reflection; team work and organisational learning. The new items introduced contribute to an assessment of to what degree teacher education contributes to teachers’ professional development and of competence to participate in school development. Since we have not been able to obtain similar assessment instrument in teacher education programs in Norway, we consider the reported instrument as a unique contribution to the scientific dialogue on teacher education in this country.

8. Limitations of the Study

First, we need to emphasise that our sample is drawn from teacher education programs at universities in Norway that prepare students for teaching grades 8 - 13. Further research and assessment of the instrument should be done for programs that prepare students to teach grades 1 - 10. Second, we acknowledge that the results are strictly limited to the sample in the article. However, we argue that it is fairly large and we clearly expect that quite similar results will be obtained in other samples of teacher education students at Norwegian universities. Third, one should also investigate if the teaching dimensions we have described on the basis of students’ self-efficacy correspond with teachers’ actual practice. Fourth, we realise that the model fit is only moderately good, which is a weakness of the presented model.

Acknowledgements

We thank Knut-Andreas Christophersen (University in Oslo) for support for data analysis, senior engineer Kyrre Svarva (NTNU) for data collection processing support and Depute head of department Lise Vikan Sandvik (NTNU) for initiating and supporting for the project, and administrative support from Trine Holter PLU, NTNU.

References


Appendix

Instrument for assessment of teacher education programs

All the items were designed with a seven point Likert Scale, with “Very poorly prepared” and “Very well prepared” as the extreme text and “Average prepared” as the mid-point text.

All “Q-numbers refer to STEP (Darling-Hammond 2007) Appendix page 136/7.

How well has the teacher education program prepared you to perform these tasks as a teacher?

DH: (2006) How well do you think your teacher preparation prepared you to?

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Std. deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. To develop and plan lessons</td>
<td></td>
<td></td>
</tr>
<tr>
<td>V1_Q5</td>
<td>To develop learning outcomes based on curriculum subjects and students’ experiences, interests and skills</td>
<td>4,74</td>
</tr>
<tr>
<td>V2_Q7</td>
<td>To develop interdisciplinary teaching programs alone or with others</td>
<td>3,51</td>
</tr>
<tr>
<td>V3_Q1</td>
<td>To develop students’ knowledge, skills and understanding in subjects using diverse methods</td>
<td>4,99</td>
</tr>
<tr>
<td>V4_Q25</td>
<td>Using knowledge of learning, subjects, curriculum and student development to promote students’ learning and professional development</td>
<td>4,83</td>
</tr>
<tr>
<td>V5_Q9</td>
<td>To adapt academic learning in school to students’ reality</td>
<td>4,43</td>
</tr>
<tr>
<td>V6_Q14</td>
<td>To justify the choice of methods of tuition for students, parents and colleagues</td>
<td>4,71</td>
</tr>
<tr>
<td>V7_Q6</td>
<td>To assess whether learning materials are suitable for my students</td>
<td>4,15</td>
</tr>
<tr>
<td>V8_Q18</td>
<td>To develop students’ ability to inquire and discuss</td>
<td>4,37</td>
</tr>
<tr>
<td>V9_Q8</td>
<td>To develop teaching strategies that promote student learning</td>
<td>4,78</td>
</tr>
<tr>
<td>2. To lead and support different students’ learning</td>
<td></td>
<td></td>
</tr>
<tr>
<td>V1</td>
<td>To lead learning processes in different subjects</td>
<td>4,67</td>
</tr>
<tr>
<td>V2</td>
<td>To motivate students</td>
<td>4,65</td>
</tr>
<tr>
<td>V3_Q26</td>
<td>To understand factors outside school which could influence students’ personal development and learning</td>
<td>4,63</td>
</tr>
<tr>
<td>V4_Q21</td>
<td>To teach students from a multicultural perspective</td>
<td>3,58</td>
</tr>
<tr>
<td>V5_Q24</td>
<td>To encourage students to interpret and to ask questions from different perspectives.</td>
<td>4,11</td>
</tr>
<tr>
<td>V6_Q10</td>
<td>To understand how students’ social, emotional and cognitive development helps learning</td>
<td>4,63</td>
</tr>
<tr>
<td>V7</td>
<td>To engage students in collaboration with others and in individual learning activities</td>
<td>4,67</td>
</tr>
<tr>
<td>V8_Q19</td>
<td>To understand how different students learn</td>
<td>4,71</td>
</tr>
<tr>
<td>V9</td>
<td>To design teaching that promotes students’ social development</td>
<td>4,19</td>
</tr>
<tr>
<td>V10</td>
<td>To engage student participation</td>
<td>4,36</td>
</tr>
<tr>
<td>3. To use assessment to guide and support for learning</td>
<td></td>
<td></td>
</tr>
<tr>
<td>V1_Q29</td>
<td>Give feedback to students to guide their learning</td>
<td>4,95</td>
</tr>
<tr>
<td>V2_Q30</td>
<td>Helping students to assess their own learning</td>
<td>4,18</td>
</tr>
<tr>
<td>V3</td>
<td>Giving constructive feedback to students to promote social development</td>
<td>4,18</td>
</tr>
<tr>
<td>V4_Q27</td>
<td>Collaborating with parents to understand students’ needs and to support their learning</td>
<td>2,62</td>
</tr>
<tr>
<td>V5_Q28</td>
<td>Using a range of assessment methods to determine student’s competence and skills</td>
<td>4,63</td>
</tr>
<tr>
<td>V6</td>
<td>Using student feedback as a basis for adapting teaching and learning</td>
<td>4,33</td>
</tr>
<tr>
<td>V7</td>
<td>Involve learners in setting their own goals for learning</td>
<td>4,07</td>
</tr>
<tr>
<td>4. To create a good learning environment (Not used: DH: Q34, Q4, Q3, Q15).</td>
<td></td>
<td></td>
</tr>
<tr>
<td>V1</td>
<td>To create a good and trusting learning environment</td>
<td>4,81</td>
</tr>
<tr>
<td>V2</td>
<td>To support students so they can achieve good results</td>
<td>4,79</td>
</tr>
</tbody>
</table>
### Continued

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>V3</td>
<td>To provide challenges and to adapt your expectations to students’ learning and achievements</td>
<td>4,57</td>
<td>1,19</td>
</tr>
<tr>
<td>V4</td>
<td>To be able to tackle the most disruptive students</td>
<td>2,81</td>
<td>1,51</td>
</tr>
<tr>
<td>V5</td>
<td>Persuading students to follow school rules.</td>
<td>3,32</td>
<td>1,45</td>
</tr>
<tr>
<td>V6_Q3</td>
<td>To help students to work independently</td>
<td>4,22</td>
<td>1,26</td>
</tr>
<tr>
<td>V7_Q12</td>
<td>Teaching in ways that support students with a mother tongue other than Norwegian</td>
<td>2,43</td>
<td>1,46</td>
</tr>
<tr>
<td>V8_Q20</td>
<td>Using good verbal and nonverbal communication strategies to guide students’ learning and behavior</td>
<td>3,87</td>
<td>1,45</td>
</tr>
</tbody>
</table>

5. Working with teacher professional development

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>V1_Q36</td>
<td>Taking leadership responsibility at different levels in school</td>
<td>3,73</td>
<td>1,62</td>
</tr>
<tr>
<td>V2_</td>
<td>To research my own practice</td>
<td>4,90</td>
<td>1,47</td>
</tr>
<tr>
<td>V3_Q35</td>
<td>To collaborate and solve problems together with colleagues.</td>
<td>4,78</td>
<td>1,41</td>
</tr>
<tr>
<td>V4_</td>
<td>Collaborating with colleagues to develop teaching</td>
<td>4,78</td>
<td>1,43</td>
</tr>
<tr>
<td>V5_</td>
<td>Collaborate to develop in learning the collegium</td>
<td>4,11</td>
<td>1,44</td>
</tr>
<tr>
<td>V6_Q16</td>
<td>Using new technologies in teaching.</td>
<td>4,37</td>
<td>1,62</td>
</tr>
<tr>
<td>V7_Q33</td>
<td>To resolve interpersonal conflicts</td>
<td>3,57</td>
<td>1,42</td>
</tr>
<tr>
<td>V8_</td>
<td>Working with continuous professional development</td>
<td>4,31</td>
<td>1,37</td>
</tr>
<tr>
<td>V9_</td>
<td>To perform research and developmental work</td>
<td>4,62</td>
<td>1,60</td>
</tr>
<tr>
<td>V10</td>
<td>To participate in improvement efforts with colleagues based on knowledge of their own school</td>
<td>3,96</td>
<td>1,54</td>
</tr>
<tr>
<td>V11</td>
<td>To reflect critically on academic, professional-ethical and educational policy issues.</td>
<td>5,03</td>
<td>1,43</td>
</tr>
<tr>
<td>V12</td>
<td>To be able to learn from others</td>
<td>5,23</td>
<td>1,41</td>
</tr>
</tbody>
</table>

Submit or recommend next manuscript to SCIRP and we will provide best service for you:

Accepting pre-submission inquiries through Email, Facebook, LinkedIn, Twitter, etc.
A wide selection of journals (inclusive of 9 subjects, more than 200 journals)
Providing 24-hour high-quality service
User-friendly online submission system
Fair and swift peer-review system
Efficient typesetting and proofreading procedure
Display of the result of downloads and visits, as well as the number of cited articles
Maximum dissemination of your research work
Submit your manuscript at: [http://papersubmission.scirp.org/](http://papersubmission.scirp.org/)