

**This is an electronic reprint of the original article.
This reprint *may differ* from the original in pagination and typographic detail.**

Author(s): Saloviita, Timo; Tolvanen, Asko

Title: Outcomes of primary teacher education in Finland : an exit survey

Year: 2017

Version:

Please cite the original version:

Saloviita, T., & Tolvanen, A. (2017). Outcomes of primary teacher education in Finland : an exit survey. *Teaching Education*, 28(2), 211-225.
<https://doi.org/10.1080/10476210.2016.1245281>

All material supplied via JYX is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the repository collections is not permitted, except that material may be duplicated by you for your research use or educational purposes in electronic or print form. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone who is not an authorised user.

Outcomes of primary teacher education in Finland: An exit survey

Timo Saloviita

Department of Teacher Education, University of Jyväskylä, Jyväskylä, Finland

Asko Tolvanen

Methodology Centre for Human Sciences, University of Jyväskylä, Jyväskylä, Finland

Corresponding author: Timo Saloviita, Department of Teacher Education, University of

Jyväskylä, PO Box 35, 40014 Jyväskylän yliopisto, Finland. Phone: +358

408053373, E-mail: timo.saloviita@jyu.fi

Asko Tolvanen, Methodology Centre for Human Sciences, University of Jyväskylä, PO

Box 35, 40014 Jyväskylän yliopisto, Finland. Phone: +358 [408053537](tel:+358408053537), E-mail:

asko.j.tolvanen@jyu.fi

Word count: 7579

Outcomes of primary teacher education in Finland: an exit survey

Seven final year cohorts of Finnish pre-service primary teachers (N = 384) were given an exit survey, which measured their estimated attainment of knowledge and experience in the 10 domains of professional activity considered critical for new teachers in the INTASC standards. The results indicated that the profile of the graduates was uneven, showing low levels of satisfaction in the fields of special education and cooperation but high levels of satisfaction in the fields of reflective practice and planning. The satisfaction with the estimated knowledge and experience achieved was divided into three components of pedagogic, instructional, and complex skills satisfaction. The central role of teaching practice was considered to be reflected in the results. The areas of low satisfaction were those in which the amount of teaching practice in the teacher education curriculum was considered insufficient to generate enough experience. This was especially so in the field of special education and in meeting the diverse needs of students. The 10-item INTASC scale showed sufficient psychometric qualities. The results are also discussed in terms of the suggested exceptional quality of Finnish teacher education.

Keywords: pre-service teachers; survey; teacher education outcomes; Finland

Introduction

Interest in evaluating the efficacy of teacher education programs is increasing (Darling-Hammond, Newton, & Chung Wei, 2010). Such evaluation can help program organisers to identify any weak points and motivate improvements. Some studies have shown that there are differences in teacher education program's outcomes, even within a single country (Wayne & Youngs, 2003). However, the task of evaluation is complicated due to the need for valid and reliable outcome measures. Also, because of the large number of interacting factors, it may be difficult to prove the existence of causal relationships between the aspects of teacher education and the outcome measures.

The evaluation of the efficacy of teacher education programs can be conducted in several ways. Input measures have included, for example, the number of credits taken, content knowledge examinations, or credit point averages. Output measures have included the use of learning outcome data, standardised observations, and surveys (Caughlan & Jiang, 2014; Worrell et al., 2014). The use of tests that measure content knowledge and pedagogical skills when making licensure decisions has been widely applied in the USA (Youngs, Odden, & Porter, 2003). Even if these tests have been criticised due to insufficient validity and reliability, strong relationships have been found between some licensure tests and student achievement (Hanushek, 1992). On the other hand, some studies have found no association between teachers' licensure scores and classroom student achievement (Buddin & Zamarro, 2008).

Measures that directly focus on outcomes, such as new teachers' effect on student achievement, are the most sophisticated type of data to be used for program evaluation. Unfortunately, this type of data is also the most difficult to collect. One example of this type of research is an Italian study, in which the scholastic

achievements of the students of over 2000 teachers were collected for two consecutive scholastic years (Caprara, Barbanelli, Steca, & Malone, 2006). The final examination grades were used as a measure of students' achievement, and previous levels of achievement were controlled. This kind of control procedure, in which students' test score gains are measured, is a variety of the "value-added model" and is promoted as a means of trustworthy teacher evaluation (Darling-Hammond, Eiler, & Marcus, 2002). However, organising this kind of measurement is generally beyond the resources of teacher education programs and so would only be possible if partnerships could be established between a large numbers of schools.

Another means of evaluating teacher education programs is direct observation in classrooms, which can identify whether the teachers apply effective practices (Pianta & Hamre, 2009). Standardised classroom observational measures have been developed for this purpose, although the problem with such an approach is again one of cost (Cameron, Connor, & Morrison, 2005; Caughlan & Jiang, 2014; Goe, Bell, & Little, 2008; Youngs, Odden, & Porter, 2003).

In contrast to observations and value-added test data, surveys using professional standards as yardsticks can provide data on a large number of participants at a low cost. Such surveys can be directed to employers, graduate students, or graduating teachers themselves. Surveys distributed to teachers typically assess their personal ratings of the teacher education program they have attended, as well as soliciting their appraisal of how well prepared they felt at the completion of their studies. This approach was applied by Thomas and Loadman (2001), whose survey contained 65 items measuring the preparation program's efficacy in four fields: teacher education graduates' careers, quality of program, teaching skills, and teaching knowledge satisfaction. The responses

given by the graduates were compared with national survey data. The results indicated that the university under study surpassed national norms.

Large international comparisons have been made of student outcomes, such as the Program for International Student Assessment (PISA) by the Organisation for Economic Co-operation and Development (OECD), which focused on 15-year-old students, or the Trends in International Mathematics and Science Study (TIMSS) assessment, which focused on students' mathematical and science knowledge. When students of a particular country exhibit high-level scores, this is sometimes attributed to a correlating high quality of teacher education (Darling-Hammond & Rothman, 2011; Sahlberg, 2011; Tryggvason, 2009). However, such claims have remained purely speculative in the absence of solid research.

While Finland has maintained a strong standing in international student comparisons, there are some data indicating lower success rates regarding the outcomes of Finnish teacher education programs. A large-scale Teaching and Learning International Survey (TALIS) by the OECD showed that Finnish lower-level secondary education teachers scored the lowest or nearly the lowest in a comparison between 32 countries regarding their felt preparedness for the content, pedagogy and practice of teaching (OECD, 2014a; OECD, 2014b, p. 501).

INTASC model core teaching standards

To properly evaluate teacher training programs, it is important to know which kinds of competencies are needed for successful professional performance. A proposal to create such a knowledge base was developed in the USA by the National Board for Professional Teaching Standards (NBPTS) to guide the assessment of teachers (Darling-Hammond, Amrein-Beardsley, Haertel, & Rothstein, 2012). On the basis of these standards, in 1992 the Interstate New Teacher Assessment and Support

Consortium (INTASC), operating within the Council of Chief State School Officers (CCSSO), developed 10 core competences for new teachers (Yinger, & Hendricks-Lee, 1998; INTASC, 2007). These INTASC standards were developed by personnel from 17 state education agencies and representatives of the teaching profession (Kim, Andrews, & Carr, 2004). The INTASC standards have undergone some reorganisation, so that, in the most recent version, the contents concerning the standards of motivation and management have been divided into other standards and a new standard, the application of content, has been added (INTASC, 2013). The 10 original core standards used in the present study are listed in this study. The standards are now intended to outline the foundations of good teaching practice for all basic education teachers across all grade levels and subject areas (INTASC, 2013). In the manual, each standard is further divided into several individual items which measure a specific field. The INTASC standards have been used to write survey items for evaluative purposes and they have become popular for teacher assessments in the USA (Darling-Hammond et al., 2012; Wilkerson, 2012).

Teacher education in Finland

Primary school teachers in Finland work in grades 1–6, having pupils between 7 and 12 years. Since 1979, primary teacher education has been based on a master's degree which takes five years to attain. In 2005, Finnish teacher education was converted to a 3 + 2 year model, whilst teacher competence was attained only after completing the whole five-year program. The master's degree involves a total of 300 credit points measured as ECTS (European Credit Transfer System) units, which is the standard for comparing the academic attainment of university students across the European Union. One credit point equals about 27 hours of study. According to the legislation (Government of Finland, 1998), the master's degree must include at least 60 credit points (ECTS) of

multidisciplinary studies on school subjects, such as mother tongue, mathematics, arts and crafts etc., and at least 60 credit points of teachers' pedagogic studies.

Typically, a primary teacher's studies include education as a major (about 140 ECTS), and 60 ECTS of this major is defined as "teachers' pedagogic studies". Education as a major involves study modules on didactics, child psychology, educational philosophy, educational sociology, and research methodology. Thesis writing and modules on scientific methodology account for almost half of the contents of a major. The major also includes teaching practice of about 20 to 25 ECTS dispersed over several years. Other parts of the studies include multidisciplinary studies on school subjects (60 ECTS), and a freely chosen minor, which is generally a school subject (60 ECTS). The remaining 40 ECTS are made up of orienting studies, like languages, and elective studies. University courses in education are roughly divided in two parts: teaching practice conducted mainly in the state-owned "normal schools" affiliated with the university on one hand, and more theoretically-oriented studies performed in the university premises, on the other hand.

The success of 15-year-old Finnish basic education students in international comparisons, such as PISA of the OECD, has strengthened the Finnish school system's international reputation (OECD, 2014c). These good results have been attributed to several factors, including the uniformity of the basic education system, highly competent teachers and the autonomy given to schools (Ministry of Education, 2016). Finnish teacher education in particular has been praised for its exceptionally high quality (Darling-Hammond & Rothman, 2011; Sahlberg, 2011), with an emphasis on its focus on a wide scientific knowledge base, research orientation and integration of theory and practice (Krokfors et al., 2011; Sahlberg, 2011). Still, the proposed reasons for this success have remained somewhat speculative. In several other countries, similar or

almost similar features exist. For prospective primary teachers, a master's-level degree was required in nine European countries in 2012 (EACEA P9 Eurydice, 2012, p. 14) and in at least 12 OECD countries in 2014 (OECD, 2014b, p.502). Also, a third of European countries grant a high degree of local autonomy to schools (EACEA P9 Eurydice, 2012, p. 13). With regard to teacher education, the requirements for entry into primary teacher training differ across OECD countries, the minimum requirement typically being an upper secondary diploma (OECD, 2014b, p. 498). Also, primary pre-service teachers must typically receive courses in pedagogy, academic subjects and educational psychology and participate in a teaching practicum, which is typically organised according to the concurrent model, so that practical training is provided at the same time as courses in specific subject matter (OECD, 2014b, p. 497). Research skills are also often required. The teaching profession in Finland is also very popular—so much so, in fact, that only about one out of every 10 applicants will be accepted into primary teacher education programs. Despite this, however, convincing evidence is lacking for the claim that “only Finland’s best and brightest” are accepted into teacher education programs (Sahlberg, 2011). The validity of entry criteria and the reliability of selection processes may reasonably be questioned due to the lack of research in this area.

In Finland, teacher education was transferred to universities in the 70s and it was elevated to master's level at the end of the same decade, even though this was met, at first, with some opposition (Niemi, 2005). International evaluation in 1994 estimated the general level of this Finnish research-based teacher education as high, but observed problems in the combination of theory and praxis, both on the level of teaching practice and research studies (Niemi, 2005). Another international evaluation estimated that the pedagogical practices in comprehensive schools were conservative and teacher-centred: “whole classes following line by line what is written in the textbooks, at a pace

determined by the teacher” (Norris, Asplund, MacDonald, Schostak, & Zamorski, 1996, p. 85). In effect, several domestic evaluation reports criticized teacher education for the insufficient investment in and focus on child diversity and professional cooperation (Niemi, 2005). It was also argued that the teacher studies remained on a level that is considered too theoretical.

Aims of the study

The aim of the present study is to survey the satisfaction of graduating pre-service teachers with the perceived competencies attained during their teacher training in one Finnish university, the University of Jyväskylä. This university runs the most established teacher education program in Finland (the course dates from 1863). The 10 essential knowledge domains detailed in the INTASC standards were used as the basis for the survey questions and the graduating students were asked the extent to which they had gained knowledge and experience in each specified field during their studies.

“Knowledge” was mainly thought to be gained from the university lectures, while “experience” was expected to originate from the teaching practice in normal schools.

These aspects were not separated, but were instead asked in combination. Comparisons were made across competency domains and across several age cohorts. In order to measure the psychometric qualities of the evaluation scale, its structure was explored using principal component analysis, and several other measures of the reliability and validity were used. In order to study the convergent and divergent validity of the INTASC scale, two other constructs were measured in one age cohort each. The selected dimensions were teacher efficacy for inclusive practices and teacher attitudes towards inclusive education. It was hypothesised that the INTASC scale would correlate high with the first and low with the second dimension.

Methods

Participants and data collection

The participants in this study were pre-service teachers who had just completed their studies at the University of Jyväskylä. The questionnaire data was collected individually from every graduating student from 2008 to 2014 on the occasion of his or her last study requirement, the final examinations. The questionnaire along with a return envelope was handed out at the final examination. In the covering letter, it was stressed that answering the questionnaire was based on the free choice of the participant, was anonymous and had no effect on the grade of the student. The number of returned questionnaires was 384 amounting to a return rate of 65%. Of the participants, 85 were men and 299 women. The age of the participants varied between 23 and 53 with a mean of 28 years. The majority, or 66%, were 25 to 28 years old. The return rates and other demographic data are presented in Table 1.

[Insert Table 1 near here]

Questionnaire

The questionnaire included questions on the date of birth and gender of the participants. It also contained an open question in which the participants could express their opinions on their study program as a whole.

INTASC scale of satisfaction

In order to measure the perceived sufficiency of knowledge and experience in the different competence domains covered during teacher education, a ten-item list of core competences for teachers was used. The 10 INTASC standards (INTASC, 2007) were transformed into a scale to measure satisfaction by adding to each of the 10 domains the introductory phrase “In my studies I have gained enough knowledge and experience of...”. Each item was measured using a Likert scale ranging from strongly disagree (1)

to strongly agree (5). The responses to the 10 statements were totalled to form a sum scale. The items and their phrasings are presented in Table 2.

The teacher efficacy for inclusive practices (TEIP)

In order to study aspects of the validity of the INTASC scale, other scales were used. The TEIP scale (Sharma, Loreman, & Forlin, 2012) was used for one study year (2014) with 53 participants. The scale contained 18 items measured with a Likert scale assessing the perceived efficacy of the teacher to teach in inclusive classrooms. The reliability of the scale was $\alpha = .83$ in the present study.

Teacher attitudes toward inclusive education scale (TAIS)

Teacher attitudes toward inclusion were measured during 2014 with 53 participants. A ten-item scale (TAIS) employing a Likert scale assessment was used. The reliability of the scale was $\alpha = .89$ and its unidimensionality was supported by factor analysis (Saloviita, 2015).

Statistical analysis

The data was analysed using IBM SPSS Statistics version 22. Some key results were presented as means and percentages. Statistical tests were conducted and Cohen's *d* was used to present the effect sizes. Principal component analysis was performed in order to study the structure of the INTASC scale. Correlation coefficients were used for the analysis of convergent and divergent validity. Polynomial contrast via a one-way ANOVA was used to test the linear trend across years in the INTASC scale sum and principal component scores.

[Insert Table 2 near here]

Results

Properties of the scale

Principal component analysis was performed for the 10 items of the INTASC scale. Three components exhibited an eigenvalue greater than 1. They explained 46.4% of the variance of the original variables. The applicability of the analysis for the data was confirmed by the Kaiser-Meyer-Olkin Measure of Sampling Adequacy (.679), indicating a mediocre acceptability, and by Bartlett's test of sphericity ($p < .000$). Since the oblique rotations indicated that the intercorrelations between the extracted components remained low, an orthogonal varimax rotation was chosen.

The varimax rotation produced three components with the first component loading high on motivation, reflexivity, communication, and child psychology (Table 3). This first component was termed Pedagogic Skills and explained 16.3% of the variance. The second component loaded high on planning, assessment, content pedagogy, and instruction, and was named Instructional Skills. It explained 15.1% of the variance. The third component was designated a component of Complex Skills and it explained 14.9% of the total variance. It loaded highest on cooperation and special education, two areas which are closely interconnected because of the need for collaborative teamwork and because the involvement of children's families is regularly recorded as an essential element of quality special education (European Agency for Development in Special Needs Education, 2009; Giangreco, 1997). Also, these two domains attained the lowest means in the raw scores, indicating that some extra difficulties were associated with them in teacher education.

Standardised component scores were calculated for further analysis using the regression method. None of these three components had a statistically significant correlation with the TAIS scale. Instead, they all had a positive correlation with the TEIP scale: Instructional Skills ($r = .467^{**}$), Complex Skills ($r = .297^{*}$) and Pedagogic Skills ($r = .283^{*}$).

The sum scores of the INTASC scale could have any value between 10 and 50. The obtained scores varied between 19–46. The observed mean of the distribution was $M = 35.01$, and standard deviation $SD = 4.46$. The value of skewness was $-.125$ with a standard error of $.125$, and kurtosis $-.092$ indicating that the distribution was close to a normal distribution. The reliability of the INTASC scale was $\alpha = 0.61$, the item-total correlations being between $r = .18$ and $.38$. The sum total of the scores over the years did not change significantly (Table 4). The correlation of the INTASC scale with the TAIS was $r = .146$, $p = .296$, indicating sufficient divergent validity, and with the TEIP $r = .557$, $p < .000$, supporting the convergent validity of the scale.

[Insert Table 3 near here]

Outcomes

The mean scores across all 10 items and all study years are presented in Table 4. Most of the means were near the middle point or number three on the scale. During all seven study years, the highest means were achieved in the domains of Planning and Reflexive Practice. The lowest means every time were achieved in the domain of Special Education, followed by Cooperation and Assessment. Of special interest is the proportion of pre-service teachers who, at the exit phase of their studies, considered their knowledge and experience insufficient. Table 2 presents the percentage of participants who disagreed or agreed with specific statements. In Table 2, the response alternatives, ‘agree’ and ‘strongly agree’ were counted together, as well as the alternatives: ‘disagree’ and ‘strongly disagree’.

When all the study years were compared domain by domain, a statistically significant difference was found in two domains, namely Assessment and Special Education (Table 3). Post hoc (Tukey) tests indicated that in the domain of Assessment the final year, 2014, differed from the first four measurement years. In the domain of

Special Education, the years 2008 and 2013 differed from each other. In the sum total, no statistically significant differences were observed between the study years.

Polynomial contrast via one-way ANOVA was used to test the linear trend across years. For the sum scores, a positive linear trend was statistically significant (contrast estimate = 1.204, standard error = 0.604, $p = 0.047$). From the three principal components, a statistically significant linear trend at the level of $p < 0.05$ appeared only in the component of Complex Skills (estimate = 0.321, standard error = 0.135, $p = 0.018$), whilst in the component of Instructional Skills, the trend was significant only at the level of $p < 0.10$ (estimate = 0.248, standard error = 0.135, $p = 0.067$). In the Pedagogic Skills component, no changes were observed (estimate = -0.086, standard error = 0.135, $p = 0.523$).

[Insert Table 4 near here]

Two differences were found between men and women. Men achieved higher scores than women in Special Education, $t(382) = -2.43$, $p = .016$, indicating a small effect with Cohen's $d = -.30$. Men achieved also higher scores in Content Pedagogy, $t(146,404) = -2.29$, $p = .030$, indicating a small effect with Cohen's $d = -.27$. When the standardised component variables were compared, no differences between men and women were noted. However, a low positive correlation between age and Complex Skills was observed, $r = .223$, $p < .000$. The TEIP scale measuring teacher efficacy in inclusive practices had statistically significant correlations with all three components, having the highest correlation with Instructional Skills, $r = .501$, $p < .000$, and lower correlations with Complex Skills, $r = .379$, $p = .005$, and Pedagogic Skills, $r = .300$, $p = .029$.

The open question was analysed by categorising the texts on the basis of their contents into subdomains. The replies of a total of 146 students from the population of

225 students from the response years 2008-2011 contained written comments for this item. The most frequently expressed wish was to gain more knowledge on the education of children with diverse needs (n = 35) and to undertake more teaching practice (n = 35), especially outside the Normal School, which was considered too artificial an environment. Several students also wanted more knowledge on cooperation with families (n = 19) and on content pedagogy (n = 15). Also, the multidisciplinary study modules were criticised for being too fragmentary and superficial (n = 36).

Discussion

Even though the present study was conducted at only one of the eight Finnish teacher preparation universities, the generalizability of its results is good due to the high homogeneity of all the teacher education programs throughout the country, which are based on a detailed legislation concerning their scope and content (Act 794/2004, 2004).

The graduating pre-service primary teachers' beliefs were examined over five years from 2008 to 2014 by means of an exit survey. The participants were asked to evaluate whether or not they had received enough knowledge and experience during their teacher studies in 10 competency domains obtained from the INTASC standards. The psychometric properties of the ten-item scale were considered sufficient, even if the reliability of the scale was just above the acceptable level. Three components emerged from the principal component analysis: Pedagogic Skills, Instrumental Skills, and Complex Skills.

Of the first two components, Pedagogic Skills and Instrumental Skills, the first seems to be more strongly associated with the child and the second with the curriculum. The component structure replicates the famous division already made by John Dewey (1902), who argued that the opposition of the child and the curriculum lay, 'below all other divisions in pedagogic opinion' (p. 4). The interest in 'teaching the child' is

expressed in the observations concerning the individual needs of the child, whilst the interest in 'teaching the subject' manifests itself in the elaborations about content knowledge of the subjects studied. As Dewey put it, 'one school fixes its attention upon the importance of the subject-matter of the curriculum as compared with the contents of the child's own experience', whilst for the others, 'the child is the starting-point, the center [sic], and the end' (pp. 7-9). The third component, Complex Skills, expresses complicated skills not easily achieved in teacher education, which are associated with meeting diversity in the classroom.

The overall mean satisfaction score across all study years and items was 35.01 points, while the midpoint of the scale distribution was 30. There was high stability in the mean scores of the domains and sum total scores over all study years. The results show that the graduating students gave positive ratings somewhat more often than negative ratings regarding the sufficiency of their studies, but their satisfaction was not overwhelming. In half of the inquired domains, a quarter of the participants disagreed, and in one domain, the majority of the participants disagreed. Only in three domains did more than 75% of the participants agree on the sufficiency of their studies.

However, it should be noted that caution is important when interpreting these data. The results of survey research are highly dependent on the exact wording of the items (Fowler, 2014). Studies asking about customer satisfaction usually give high satisfaction percentages, no matter what the actual level of service has been. Thus, in the meta-analysis of 221 studies on medical care, the general level of customer satisfaction varied from 76% to 84% (Hall & Dornan, 1988). Only when the questions were more specific, the levels of satisfaction tended to drop (Williams & Calnan, 1991). Because an absolute level of satisfaction may at least partly be an artefact of the survey

method itself, it is more interesting to look at the differences, connections and trends in percentages than their absolute levels.

A small linear trend in sum total scores was observed indicating that total scores turned more positive during the years. More specifically, this increase was observed in the component of Complex Skills and less conspicuously in the component of Instructional Skills. This positive trend possibly reflects changes realised in the curriculum which might lead to a slightly greater student satisfaction. However, it is not possible to track their true origins. These positive changes in student evaluations may be a result of restructuring study modules, or they may reflect, for example, a staff reshuffle.

Of the three components, Instructional Skills explained 22% of the variance of the scores of the Teacher Efficacy for Inclusive Practices Scale (TEIP). Other components had lower explanation rates with this scale: Complex Skills 9% and Pedagogic Skills 8%. These results indicated that teacher self-efficacy in managing diversity in the classroom was best explained by their general instructional skills.

The participants gave very different evaluations of their competencies across various domains. The majority of participants felt that they had not gained enough knowledge and experience in the field of special education or diversity in the classroom, while almost half regarded their competencies in school and community cooperation to be insufficient. This result was also reflected in the open question, in which almost a quarter of participants wanted more knowledge on special education. This finding was in accordance with several teacher education evaluation reports that have indicated similar shortages in teacher education (Ministry of Education, 2007). On the other hand, in terms of the planning and reflective practice domains, very few participants lamented gaining too little knowledge and experience. Comparative results were obtained in the

large study on lower secondary education teachers who reported special education as the primary field in need of professional development, whilst knowledge of the curriculum was the least mentioned topic (OECD, 2014d).

These differences probably reflect some of the key problems with teacher training. Pre-service teachers have only scant opportunity to acquire practical experience of diverse students during their teaching practice. One reason for this may be the relative lack of children with special needs in Finnish mainstream classrooms. According to statistics, about 5% of all children in basic schools are educated in separate special education classrooms, which is one of the largest numbers reported in Europe (European Agency for Development in Special Needs Education 2012a; Statistics Finland, 2014). Also, the short duration of the teaching practice modules may explain the result. Pre-service teachers typically have insufficient time to become absorbed in the problems of individual students or to engage in cooperation with families or other actors in the school. Fuller and Bown (1975) describe the pre-service teachers' concerns in the classrooms as initially deriving from personal survival concerns and progressing to teaching situation concerns. Only after this is the student teacher able to direct attention to pupil-related concerns. It may be that the relatively short length of teaching practice does not suffice to provide enough opportunities for student teachers to arrive at this final phase in which the needs of individual pupils are recognised. However, every lesson must be planned and, therefore, it is no surprise that Planning was a highly scored competency.

These considerations seem to stress the importance of teaching practice as a crucial element in teachers' education when considering the attainment of all necessary teacher competencies. Despite its significance, the extent of teaching practice was limited in the study university (about 25 credit points out of 300), probably because it is

the most expensive part of teacher education. The question arises as to whether teaching practice should be enhanced through increasing the hours dedicated to it. It might also be fruitful to build closer connections between teaching practice and more theoretical modules in pre-service teacher education, thus filling the gap between theory and practice as recommended in the evaluation reports (Niemi, 2005). The results of the open question also stressed the importance of teaching practice. Almost a quarter of students wanted more teaching practice experience in their program.

The low level of satisfaction in the fields of Special Education and Cooperation needs attention because it may hinder the development of inclusive education, a goal confirmed by the United Nations Convention on the Rights of People with Disabilities (2006). The European Agency for Development in Special Needs Education (2012b) identified three key areas of competence that relate to the ability to teach diverse students: support for all learners, cooperation, and personal professional development. In the present study, two of these three areas were domains that scored the lowest.

It has been shown that teachers' sense of efficacy underlies their placement preferences regarding the children with special education needs. Primary teachers with a higher level of self-efficacy were more likely to agree to inclusive placements than those teachers with lower efficacy (Soodak & Podell, 1993). The observed low level of satisfaction with their competence to teach children with special education needs among the graduates may therefore result in teachers who continue to prefer segregated special education placements in their work. In order to enhance the abilities of future teachers in this respect, a pivotal issue may be the possibility for everybody to undertake teaching practice in inclusive classrooms and thus strengthen their skills in meeting diverse student needs.

The age and sex of the participants had some connections with the INTASC domains. The increased age of participants had a small positive correlation with the Complex Skills component, which possibly indicates that older graduates had acquired more real-life teaching experience than the younger graduates. Men had higher scores than women in Content Pedagogy and Special Education. Since both genders completed exactly the same studies, these differences may reflect gender role differences between male and female students.

The shortness of the scale used to measure the teacher training outcomes can be regarded its limitation. The shortness was, however, intentional because short scales, if psychometrically sound enough, provide an economic means to conduct surveys, which typically suffer from low return rates due to being perceived as overly long.

The item formulations in this study brought the measured dimension near to the concept of self-efficacy. According to Bandura (1997), self-efficacy refers to “beliefs in one’s capabilities to organize and execute the courses of action required to produce given attainments” (p. 3). Teachers’ self-efficacy has accordingly been understood as the confidence that teachers have in their individual and collective capability to influence students’ learning (Klassen, Tze, Betts, & Gordon, 2011). The item formulations in this study expressed the participants’ satisfaction or strength of belief in the sufficiency of their knowledge and experience, but did not specify the action. What was lacking, then, was the affirmation of the capability to carry out the action itself (see Bandura, 1997, p. 382). Future studies should consider developing efficacy measures that would be suitable for graduating pre-service teachers.

A critical question is whether teacher self-evaluations, such as the INTASC survey implemented in this study, have any predictive validity for actual teacher practice. The question is justified because surveys are based on perception, not reality

(Coggshall, Bivona, & Reschly, 2012). The problem can only be answered by studies using multiple performance indicators. In the present study, it was shown that the graduating teachers' evaluations of their level of knowledge and experience modestly correlated with their sense of self-efficacy. This correlation was highest in the component of Instructional Skills. Self-efficacy, on the other hand, has been shown to be associated with other outcome measures of teaching practice, such as job satisfaction or students' academic achievement (Caprara et al., 2006).

A challenge for future studies will be to identify predictors of good teacher achievement, which would be easy to assess at the exit phase of teacher education programs. The development of such predictors needs studies on the predictive validity of the exit measures applied.

The present study has showed that the satisfaction of Finnish graduating primary teachers regarding their studies is in general positive, but not overwhelmingly so. The general satisfaction was at a moderate level, which is typical for satisfaction studies. The profile was also uneven across subdomains, indicating conspicuous dissatisfaction in some key areas of teacher competence. On the basis of the obtained numbers in this study, claims of the outstanding excellence of Finnish teacher education may regard as overstatements.

However, explanations have been given for the Finnish PISA success other than high-quality teacher education. Some voices have attributed this success to a wider range of social, cultural and historical factors, including the late modernization of the country (Simola, 2005). It has been shown that the cultural and demographic backgrounds of students may be more influential on their educational success than their exposure to a specific educational system (Feniger & Lefstein, 2014). Others have paid attention to the government's use of PISA results as a means to justify the central

government's policy (Rautalin & Alasuutari, 2009). In this context, the attribution of PISA success to Finnish teachers' university training may well serve this agenda. International comparisons of the satisfaction of graduating teachers regarding their studies would be welcome to assess the extent of the association of the student satisfaction with the organization and content of their studies.

Acknowledgments

Responses to the open question were analysed by Tomi Hiltunen and Jaska Kunelius in their joint master's thesis.

References

- Act 794/2004 (2004). Valtioneuvoston asetus yliopistojen tutkinnoista [Government decree on university degrees]. Finlex. Retrieved from:
<http://www.finlex.fi/fi/laki/alkup/2004/20040794>
- Bandura, A. (1997). *Self-efficacy. The exercise of Control*. New York, NY: Freeman.
- Buddin, R. & Zamarro, G. (2008). *Teacher quality, teacher licensure tests, and student achievement*. RAND Education working papers. Retrieved from:
http://130.154.3.8/content/dam/rand/pubs/working_papers/2008/RAND_WR555.pdf
- Cameron, C. E., Connor, C. M., & Morrison, F. J. (2005). Effects of variation in teacher organization on classroom functioning. *Journal of School Psychology, 43*, 61-85.
- Caprara, G. V., Barbanelli, C., Steca, P., & Malone, P. S. (2006). Teachers' self-efficacy beliefs as determinants of job satisfaction and students' academic achievement: A study at the school level. *Journal of School Psychology, 44*, 473-490.

- Caughlan, S., & Jiang, H. (2014). Observation and teacher quality: Critical analysis of observational instruments in preservice teacher performance assessment. *Journal of Teacher Education*, 65, 375–388. doi: 10.1177/0022487114541546
- Coggshall, J. G., Bivona, L., & Reschly, D. J. (2012). *Evaluating the effectiveness of teacher preparation programs for support and accountability*. Washington, DC: National Comprehensive Center for Teacher Quality.
- Darling-Hammond, L., Amrein-Beardsley, A., Haertel, E., & Rothstein, J. (2012). Evaluating teacher evaluation. *Phi Delta Kappan*, March, supplement, 8-15.
- Darling-Hammond, L., Eiler, M., & Marcus, A. (2002). Perceptions of preparation: Using survey data to assess teacher education outcomes. *Issues in Teacher Education*, 11, 65-84.
- Darling-Hammond, L., Newton, X., & Chung Wei, R. (2010). Evaluating teacher education outcomes: A study of the Stanford teacher education programme. *Journal of Education for Teaching: International Research and Pedagogy*, 36, 369-388.
- Darling-Hammond, L., & Rothman, R. (2011). Lessons learned from Finland, Ontario, and Singapore. In Darling-Hammond & R. Rothman (Eds.), *Teacher and leader effectiveness in high-performing education systems* (pp. 1- 12). Washington, DC: Alliance for Excellent Education.
- Dewey, J. (1902). *The child and the curriculum*. Chicago: The University of Chicago Press.
- EACEA P9 Eurydice (Education, Audiovisual and Culture Executive Agency) (2012). *Key data on education in Europe 2012*. doi:10.2797/77414
- European Agency for Development in Special Needs Education (2009). *Key principles for promoting quality in inclusive education. Recommendations for policy makers*.

Odense: European Agency for Development in Special Needs Education.

Retrieved from: <http://www.european-agency.org/>

European Agency for Development in Special Needs Education (2012a). *Special Needs Education Country Data*. Retrieved from:

<https://www.european-agency.org/publications/ereports/sne-country-data-2012/sne-country-data-2012>

European Agency for Development in Special Needs Education (2012b). *Teacher Education for Inclusion. Profile of Inclusive Teachers*. Odense: Author.

Feniger, Y. & Lefstein, A. (2014). How *not* to reason with PISA data: An ironic investigation. *Journal of Education Policy*, 29, 845-855. doi: 10.1080/02680939.2014.892156

Fowler, F.J. (2014). *Survey research methods. 5th Edition*. Thousand Oaks, CA: Sage.

Fuller, F. F. & Bown, O. H. (1975). Becoming a Teacher. In K. J. Ryan (Ed.), *Teacher education. The seventy-fourth yearbook of the National Society for the Study of Education. Part II* (pp. 25-52). Chicago: University of Chicago Press.

Giangreco, M.F. (1997). Key lessons learned about inclusive education: Summary of the 1996 Schonell Memorial Lecture. *International Journal of Disability, Development and Education*, 44, 193-206.

Goe, L., Bell, C., & Little, O. (2008). *Approaches to evaluating teacher effectiveness: A research synthesis*. Washington, DC: National Comprehensive Center for Teacher Quality. Retrieved from

<http://www.gtlcenter.org/sites/default/files/docs/EvaluatingTeachEffectiveness.pdf>

f

Government of Finland. (1998). Act on Teacher Qualifications (14.12.1998/986).

Finlex.

- Hall, J.A. & Dornan, M.C. (1988). Meta-analysis on satisfaction with medical care: Description of research domain and analysis of overall satisfaction levels. *Social Science Medicine*, 27, 637-644.
- Hanushek, E. A. (1992). The trade-off between child quantity and quality. *Journal of Political Economy*, 100, 84-117.
- INTASC. (2007). *The INTASC Standards*. Retrieved from http://www.ccsso.org/projects/Interstate_New_Teacher_Assessment_and_Support_Consortium/Projects/Standards_Development/
- INTASC. (2013). *INTASC Model Core for Teaching Standards and Learning Progressions for Teachers 1.0*. Retrieved from http://www.ccsso.org/Documents/2013/2013_INTASC_Learning_Progressions_for_Teachers.pdf
- Kim, M. M., Andrews, R. L., & Carr, D. L. (2004). Traditional versus integrated preservice teacher education curriculum. A case study. *Journal of Teacher Education*, 55, 341-356.
- Klassen, R. M., Tze, V. M. C., Betts, S. M., & Gordon, K. A. (2011). Teacher efficacy research 1998-2009: Signs of progress or unfulfilled promise? *Educational Psychology Review*, 23, 21-43.
- Krokfors, L., Kynäslähti, H., Stenberg, K., Toom, A., Maaranen, K., Jyrhämä, R., Byman, R., & Kansanen, P. (2011). Investigating Finnish teacher educators' views on research-based teacher education. *Teaching Education*, 22, 1–13. doi: 10.1080/10476210.2010.542559
- Ministry of Education (2007). Opettajankoulutus 2020 [Teacher Education 2020], *Opetusministeriön työryhmämuistioita ja selvityksiä 2007*, 44.

Ministry of Education (2016). *Background for Finnish PISA success*. Retrieved from:

<http://www.minedu.fi/pisa/taustaa.html?lang=en>

Niemi, H. (2005). Suomalainen opettajankoulutus valmiina jo pitkään eurooppalaiseen korkeakoulualueeseen [Finnish teacher ready for a long time for the European higher education area] In: R. Jakku-Sihvonen (Ed.) *Uudenlaisia maistereita* [New types of masters] (pp.187-218). Jyväskylä: PS-Kustannus.

Norris, N., Asplund, R., MacDonald, B., Schostak, J., & Zamorski, B. (1996). *An independent evaluation of comprehensive curriculum reform in Finland*. Helsinki: National Board of Education.

OECD (2014a). *The OECD Teaching and Learning International Survey (TALIS) 2013*

Results - Excel Figures and Tables. Table 2.4. OECD Publishing. Retrieved

from: <http://www.oecd.org/edu/school/talis-excel-figures-and-tables.htm>

OECD (2014b). *Education at a glance. OECD indicators*. Retrieved from:

<https://www.oecd.org/edu/Education-at-a-Glance-2014.pdf>

OECD (2014c). *PISA 2012 Results*. Retrieved from:

<http://www.oecd.org/pisa/key/findings/pisa-2012-results.html>

OECD (2014d). *A teachers' guide to TALIS 2013. Teaching and Learning International*

Survey, TALIS. OECD Publishing. Retrieved from:

<http://dx.doi.org/10.1787/9789264216075-en>

Pianta, R. C., & Hamre, B. K. (2009). Conceptualization, measurement, and improvement of classroom processes: Standardized observation can leverage capacity. *Educational Researcher*, 38, 109-119.

doi:10.3102/0013189X09332374

- Rautalin, M. & Alasuutari, P. (2009). The users of national PISA results by Finnish officials in central government. *Journal of Education Policy, 24*, 539-556. doi: 10.1080/02680930903131267
- Sahlberg, P. (2011). Developing effective teachers and school leaders: The case of Finland. In Darling-Hammond & R. Rothman (Eds.), *Teacher and leader effectiveness in high-performing education systems* (pp. 13-212). Washington, DC: Alliance for Excellent Education.
- Saloviita, T. (2015). Measuring pre-service teachers' attitudes towards inclusive education: Psychometric properties of the TAIS scale. *Teaching and Teacher Education, 52*, 66-72.
- Sharma, U., Loreman, T., & Forlin, C. (2012). Measuring teacher efficacy to implement inclusive practices. *Journal of Research in Special Educational Needs, 12*, 12-21.
- Simola, H. (2005). The Finnish miracle of PISA: historical and sociological remarks on teaching and teacher education. *Comparative Education, 41*, 455-470. doi: 10.1080/03050060500317810
- Soodak, L. C. & Podell, D. M. (1993). Teacher efficacy and student problem as factors in special education referral. *The Journal of Special Education, 27*, 66-81.
- Statistics Finland (2014). Erityisopetus [Special education] Retrieved from: <http://www.stat.fi/til/erop/index.html>
- Thomas, A. M., & Loadman, W. E. (2001). Evaluating teacher education programs using a national survey. *The Journal of Educational Research, 94*, 195-206. doi:10.1080/00220670109598753

- Tryggvason, M-T. (2009). Why is Finnish teacher education successful? Some goals Finnish teacher educators have for their teaching. *European Journal of Teacher Education, 32*, 369-382. doi: 10.1080/02619760903242491
- United Nations. (2006). *Convention on the rights of people with disabilities*. Retrieved from <http://www.un.org/esa/socdev/enable/index.html>
- Wayne, A. J., & Youngs, P. (2003). Teacher characteristics and student achievement gains: A review. *Review of Educational Research, 73*, 89-122.
- Wilkerson, J. R. (2012). Measurement and evaluation perspectives on scaling teacher affect with multiple measures. *The International Journal of Educational and Psychological Assessment, 9*, 165-187.
- Williams, S.J. & Calnan, M. (1991). Key determinants of consumer satisfaction with general practice. *Family Practice, 8*, 237-242.
- Worrell, F., Brabeck, M., Dwyer, C., Geisinger, K., Marx, R., Noell, G., & Pianta, R. (2014). *Assessing and evaluating teacher preparation programs*. Washington, DC: American Psychological Association.
- Yinger, R. J. & Hendricks-Lee, M. S. (1998). Professional development standards as a new context for professional development in the US. *Teachers and Teaching: Theory and Practice, 4*, 273- 298.
- Youngs, P., Odden, A., & Porter, A. (2003). State policy related to teacher licensure. *Educational Policy, 17*, 217-236. doi:10.1177/0895904803017002002