

**This is a self-archived version of an original article. This version may differ from the original in pagination and typographic details.**

**Author(s):** Goller, Michael; Ursin, Jani; Vähäsantanen, Katja; Festner, Dagmar; Harteis, Christian

**Title:** Finnish and German student teachers' motivations for choosing teaching as a career : The first application of the FIT-Choice scale in Finland

**Year:** 2019

**Version:** Accepted version (Final draft)

**Copyright:** © 2019 Elsevier Ltd.

**Rights:** CC BY-NC-ND 4.0

**Rights url:** <https://creativecommons.org/licenses/by-nc-nd/4.0/>

**Please cite the original version:**

Goller, M., Ursin, J., Vähäsantanen, K., Festner, D., & Harteis, C. (2019). Finnish and German student teachers' motivations for choosing teaching as a career : The first application of the FIT-Choice scale in Finland. *Teaching and Teacher Education*, 85, 235-248.  
<https://doi.org/10.1016/j.tate.2019.06.023>

**Finnish and German Student Teachers' Motivations  
for Choosing Teaching as a Career.  
The First Application of the FIT-Choice Scale in Finland.**

**Michael Goller** (corresponding author)

University of Bamberg  
Economics and Business Education  
Kapuzinerstraße 16  
D-96047 Bamberg  
GERMANY

University of Paderborn  
Institute of Educational Science  
Warburger Str. 100  
D-33098 Paderborn  
GERMANY  
michael.goller@uni-paderborn.de

**Jani Ursin**

University of Jyväskylä  
Finnish Institute for Educational Research  
P.O. Box 35  
FI-40014 University of Jyväskylä  
FINLAND  
jani.p.ursin@jyu.fi

**Katja Vähäsantanen**

University of Jyväskylä  
Department of Education  
P.O. Box 35  
FI-40014 University of Jyväskylä  
FINLAND  
katja.vahasantanen@jyu.fi

**Dagmar Festner**

University of Paderborn  
Institute of Educational Science  
Warburger Str. 100  
D-33098 Paderborn  
GERMANY  
dagmar.festner@uni-paderborn.de

**Christian Harteis**

University of Paderborn  
Institute of Educational Science  
Warburger Str. 100  
D-33098 Paderborn  
GERMANY  
christian.harteis@uni-paderborn.de

## **Abstract**

This study introduces and validates the Finnish version of the FIT-Choice scale and uses the scale to explore Finnish undergraduate students' motivations for choosing teaching as a career as well as their perceptions of teaching as a profession in comparison to German student teachers. The results replicated the FIT-Choice structure both in Finland and Germany. Sample comparisons revealed that motivations for teaching and perceptions of the teaching profession reflected both differences and similarities between Finland and Germany. The study offers novel information on student teachers' motivational structure.

## **Highlights:**

- RQ1. Validation of the FIT-Choice structure was conducted in Finland.
- RQ2. The FIT-Choice scale functioned quite similarly in Finland and Germany.
- RQ3. Motivations for teaching seemed to be more different than similar in these two countries.
- RQ4. Perceptions of teaching seemed to be more similar than different in these two countries.

**Keywords:** FIT-Choice, career, comparative research, motivation, teaching, beginning student teachers

## **Published as:**

Goller, M., Ursin, J., Vähäsantanen, K., Festner, D., & Harteis, C. (2019). Finnish and German student teachers' motivations for choosing teaching as a career. The first application of the FIT-Choice scale in Finland. *Teaching and Teacher Education*, 85, 235-248. <https://doi.org/10.1016/j.tate.2019.06.023>

**Finnish and German Student Teachers' Motivations  
for Choosing Teaching as a Career.  
The first Application of the FIT-Choice Scale in Finland.**

Michael Goller<sup>a,b</sup>, Jani Ursin<sup>c</sup>, Katja Vähäsantanen<sup>c</sup>, Dagmar Festner<sup>b</sup>, & Christian Harteis<sup>b</sup>

<sup>a</sup> *University of Bamberg*, <sup>b</sup> *University of Paderborn*, <sup>c</sup> *University of Jyväskylä*

**1. Introduction**

The importance of educating competent future teachers has become evident in many countries. Among others, the Programme for International Student Assessment (PISA) and the Teaching and Learning International Survey (TALIS) highlighted a crucial and often challenging role of teachers in shaping the future of societies (see also Watt, Richardson, & Smith, 2017). Furthermore, the quality of teachers is often associated with the quality of students' learning and, ultimately, the whole education system (Hattie, 2009; Sahlberg, 2011). In particular, students' motivations for choosing a teaching career have significant consequences, since such motivations seem to be relatively stable over time (Roness & Smith, 2010) and are closely connected to teaching behaviour, commitment, well-being, and students' learning (Watt et al., 2017). Moreover, career choice motivations are important predictors of teachers' progression in and graduation from teacher education programmes (König & Rothland, 2012). It is therefore an important endeavour to understand on what premise young adults actually make the decision to pursue a teaching career.

Answers to this question can be obtained using the FIT-Choice (Factors Influencing Teaching Choice) framework, since it provides a theory-based and validated scale for exploring the motivations for choosing teaching as a career and the perceptions about teaching (Watt & Richardson, 2007, 2012). So far, the framework has mainly been utilised in single country contexts and less from a cross-cultural comparative perspective. Such a comparative perspective including more than one country, however, is crucial, since the social basis of a teaching profession is created through behavioural and cultural patterns, specific artefacts, and their connection to certain social and institutional practices (e.g., Sarja, Nyman, Ito, & Jaatinen, 2017). Thus the comparative perspective allows us to understand how different cultural patterns and national practices shape the motivations of young people to become teachers. Such knowledge may later be used for governance purposes (e.g., to avoid teaching staff shortages by attracting new teachers and by reducing the turnover of already employed teachers; see also Lin et al., 2012).

In this study, we compare and explore Finnish and German students' motivation to become teachers and their more general perception about teaching as a job. Although both countries have teacher education programmes that require students to study two consecutive degrees with comparable workloads, they strongly differ in regard to entrance requirements for their student teachers, the status of the teaching profession in the general public, and the remuneration. In addition, the Finnish education system has been praised because of its good results in benchmark studies such as PISA, while Germany achieved only average scores in those assessments (OECD, 2001, 2016). Therefore, it is an interesting endeavour to find out whether student teachers from both countries differ in their career choice motivations and how such discrepancies can be explained by differences between the education systems and cultural contexts.

To our knowledge, empirical findings about Finnish students' motivations to become teachers in comparison to other countries are still lacking. One reason for this research gap might be that prior to this study no Finnish version of the FIT-Choice scale existed. It was therefore necessary to adapt the scale to the Finnish context and to empirically validate it with a Finnish sample of student teachers as part of this study.

Taken together, this study contributes to the literature threefold: First, it introduces and validates a Finnish version of the FIT-Choice scale. Second, it explores why Finnish student teachers decide to pursue teaching as a career and how they perceive teaching as a profession. Third, the results of the Finnish sample are compared to a sample of German student teachers explaining whether the findings are congruent with differences between both countries' teacher education programmes.

The paper starts with a discussion of the theoretical grounding of the FIT-Choice framework, including the empirical findings generated based on it. After this, both the Finnish and German teacher education systems are shortly described. This is followed by the description of the first part of the empirical study in which the Finnish version of the FIT-Choice scale is introduced and empirically validated. Hereafter, the findings of the second part of the empirical study comparing Finnish and German student teachers' career choice motivation are reported. The paper ends with a discussion of the empirical findings.

## **2. Conceptual framework**

### ***2.1. Various motivations for becoming a teacher***

Previous studies suggest that motivations for becoming a teacher typically relate to extrinsic, intrinsic, and altruistic reasons, the two latter ones being dominant. These motivations can be summed up as follows (Kyriacou & Coulthard, 2000; Manuel & Hughes, 2006; Papanastasiou & Papanastasiou, 1997; Roness & Smith, 2010): (a) *extrinsic motivations* cover aspects of the job that are not inherent in the work itself, such as long holidays, level of pay, and status; (b) *intrinsic motivations* cover aspects of the job activity itself, such as the activity of teaching children and having an interest in using subject matter knowledge; and (c) *altruistic motivations* involve seeing teaching as a socially worthwhile and important job, and having a desire to help children succeed as well as to help society improve.

Trying to take into account the diversity of these motivating factors of becoming a teacher, Watt and Richardson (2007, 2012) developed a multidimensional framework featuring a corresponding measurement scale called FIT-Choice founded on expectancy-value theory (Eccles et al., 1983; Wigfield & Eccles, 2000) and on a comprehensive literature review of the factors that determine why someone considers becoming a teacher. However, the FIT-Choice framework goes beyond the classical triad of extrinsic, intrinsic, and altruistic motivations by further differentiating these, as well as by specifying additional factors that also determine whether individuals choose teaching as their career (Watt & Richardson, 2007). The authors propose the following motivations for teaching: (a) *perceived teaching abilities*, (b) *intrinsic value*, (c) personal utility values (*job security, time for family, job transferability*), (d) social utility values (*shape future of children/adolescents, enhance social equity, make social contribution, work with children/adolescents*), (e) and the more negatively connoted motivation of having chosen teaching only as a '*fallback*' career, or because of (f) *prior teaching and learning experiences*, or (g) *social influences* from others.

In addition, based on expectancy-value theory, Watt and Richardson (2007) further proposed that certain perceptions regarding teaching as a job also affect the decision to become a teacher. These perceptions can be divided into perceived task demand and task return. Task demand describes the perception that teaching is *highly demanding* and requires *high levels of expertise*. Task return includes the perception that teaching is a respected job with a high *social status* and that teachers earn a comparatively good *salary*. Watt and Richardson (2007) also suggested to include information on the degree to which the student was discouraged from taking up teaching as a career (*social dissuasion*) and her/his current degree of *satisfaction with choice*. Table 1 lists and defines all factors of the FIT-Choice framework.

**Table 1.** Theoretical factor structure of the FIT-Choice framework.

<b>Factor</b>	<b>Definition</b>
Motivations for teaching	
<i>Perceived teaching abilities</i>	Subjective expectation to perform well as a teacher.
<i>Intrinsic value</i>	Interest and expected enjoyment of working as a teacher.
Personal utility values	
<i>Job security</i>	Perception that teaching protects from unemployment.
<i>Time for family</i>	Perception that teaching offers sufficient time for family members and family duties.
<i>Job transferability</i>	Perception that a teaching career allows teachers to freely choose their place of residency.
Social utility values	
<i>Shape future of children/adolescents</i>	Desire to affect children's or adolescents' prospective life trajectories and thereby to influence the upcoming generation through teaching.
<i>Enhance social equity</i>	Desire to reduce social inequality and thereby to positively affect disadvantaged students through teaching.
<i>Make social contribution</i>	Desire to make a positive contribution to society through teaching.
<i>Work with children/adolescents</i>	Desire to work in a job that mainly deals with children or adolescents.
<i>Fallback career</i>	Choosing to become a teacher because of a failure to pursue one's first-choice career.
<i>Prior teaching and learning experiences</i>	Extent to which the individual had good teachers as role models and experienced learning and teaching as something positive.
<i>Social influences</i>	Extent to which significant others (e.g., family, friends) expressed teaching as a suitable career for the individual.
Perceptions of teaching	
Task demand	
<i>High demand</i>	Perception that teaching is a highly challenging job because it poses a high workload and is also emotionally demanding.
<i>Expert career</i>	Perception that teaching requires high levels of specialised expertise.
Task return	
<i>Social status</i>	Perception that teaching is a career that is socially highly valued and respected.
<i>Salary</i>	Perception that teachers earn a comparatively good remuneration.
<i>Social dissuasion</i>	Extent to which the student was discouraged to take up teaching as a career.
<i>Satisfaction with choice</i>	Extent to which the student is satisfied with the choice of becoming a teacher.

*Note.* Definitions are mainly based on Watt and Richardson (2007). All factors set in italics are first-order constructs later measured with several different items as subscales.

The FIT-Choice measurement scale has been used and validated in various contexts: Australia (Watt & Richardson, 2007), Croatia (Jugovic, Marusic, Ivanec, & Vidovic, 2012), Germany (König & Rothland, 2012), the Netherlands (Fokkens-Bruinsma & Canrinus, 2012), Turkey (Akar, 2012; Kılınç et al., 2012), Ireland (Hennessy & Lynch, 2017), and Ghana (Salifu, Alagbela, & Ofori, 2017), as well as from a broader comparative perspective including, among others, Norway and China (Lin et al., 2012; Watt et al., 2012). More recently, the scale has also been used to investigate teacher motivations in countries like Spain and Estonia (Watt et al., 2017).

Several studies have confirmed the FIT-Choice scale structure, but some studies have also identified some differences in the structure for motivations across cultural contexts. For example, in their Irish sample, Hennessy and Lynch (2017) could not replicate all first-order factors of the scale. Instead, a few of the subscales did not exhibit discriminant validity and formed new empirical factors that had not been described by Watt and Richardson (2007). Affected were the following subscales: *work with children/adolescents* and *intrinsic values*, *job security*, and *time for family*, as well as *enhance social equity*, *shape future of children/adolescents*, and *job transferability*. In another study, Watt et al. (2012) had problems to establish sufficient reliability regarding the *job transferability* and the *fallback career* subscales in a German and a Norwegian sample. Because of these reliability issues, both of these subscales had to be removed from any further analysis in their study. Besides these issues, the factor structure of the FIT-Choice scale proved to be a reliable and valid measurement instrument to capture motivations and perceptions regarding teaching in a range of cultural settings.

## ***2.2. Factors related to the motivations for choosing teaching as a career***

The FIT-Choice framework provides a platform for comparisons across cultural contexts (Watt & Richardson, 2012). Although intrinsic and altruistic motivations seem to be the most typical ones for choosing teaching as a career in most cultures, the context also seems to matter. In Australia, where the FIT-Choice framework was originally developed, the highest-rated motivations for teaching were *perceived teaching abilities*, *intrinsic value* of teaching, the desire to *make social contribution*, as well as *shape future* and *work with children/adolescents*. The lowest rated motivations were choosing teaching as a *fallback career* or because of *social influences* in the form of encouragement from others to undertake teaching (Richardson & Watt, 2006). In the Netherlands, for example, the findings were similar, since the motivational factor of *perceived teaching abilities* was influential in the

decision to become a teacher. *Social influences, time for family, prior teaching and learning experiences*, and the aim to *enhance social equity* were the least important motivations (Fokkens-Bruinsma & Canrinus, 2012). Related research on Ireland yielded results that were partly contrary to those of the Netherlands; that is, while the motivational factor of *perceived teaching abilities* was likewise important there, the factor of *prior teaching and learning experiences* was nevertheless equally important. This finding might be specific to the Irish context because of the traditional dominance of subject-based knowledge in the Irish second-level schooling system (Hennessy & Lynch, 2017). In the German context, it seems that teachers' wish to *work with children/adolescents* dominates as the decisive factor for teaching as a career (König & Rothland, 2012; Rothland, 2014). In the United States, Turkey, and China, the participants highlighted altruistic motivations more; the respondents were motivated to enter teaching because of their *social utility values* (Kılınç et al., 2012; Lin et al., 2012). Furthermore, in Turkey, the *perceived teaching abilities* and *intrinsic value* motivations were not among the most influential ones. In the Turkish context, together with altruistic reasons, *job security* was a paramount motivation to choose teaching as a career (Kılınç et al., 2012). Consequently, the previous studies seem to suggest that socio-cultural and national contexts indeed play an important role in the motivation to become a teacher.

In terms of perceptions regarding the teaching profession, previous studies indicate that teaching as a career is often perceived as a job with relatively *low status* in almost all countries, and as offering a *low salary* in most countries except for Germany (Akar, 2012; Fokkens-Bruinsma & Canrinus, 2012; Hennessy & Lynch, 2017; Kılınç et al., 2012; König & Rothland, 2012; Lin et al., 2012; Watt & Richardson, 2007; Watt et al., 2012). Nevertheless, most student teachers reported that there was no active *social dissuasion* to pursue this career by relatives or other acquaintances. Quite a few studies also show that teaching is perceived to be a difficult job with *high demands* requiring a substantial stock of *expertise* (Akar, 2012; Hennessy & Lynch, 2017; Kılınç et al., 2012; König & Rothland, 2012; Lin et al., 2012; Watt & Richardson, 2007; Watt et al., 2012). However, Chinese pre-service teachers judged both the level of *difficulty* of the job and the required *expertise* substantially lower than teachers in other countries did (Lin et al., 2012).

### **3. Teacher education and teaching as a profession in Finland and Germany**

In this study, students from Finland and Germany were compared regarding their motivations for becoming a teacher and their perceptions of the teaching career. Both countries have somewhat different education systems, teacher education programmes, and general

perceptions of teaching as a profession, as we present next in more detail (see also Erss, 2018). However, it is worth noting that there are also some similarities between these national contexts. In particular, all of the teachers are highly qualified and a special Master's degree is required.

### ***3.1. Teacher education and teaching in Finland***

In Finland, most children younger than six years participate in non-compulsory early childhood education and care (ECEC) as well as in compulsory preschool at the age of six. Both of these are taught mostly by preschool teachers (Paronen & Lappi, 2018). After that, the children start compulsory basic school at the age of seven, which includes primary (grades 1–6) and secondary (grades 7–9) school phases. Primary school teachers teach all subjects in grades 1–6 and may also provide pre-primary education. In comparison, subject teachers teach one or several subjects in basic education, primarily in grades 7–9. Next, most adolescents enter either general upper secondary education taught by subject teachers or vocational upper secondary education taught by vocational teachers.

In Finland, teacher education takes place in the university context. All primary and subject teachers are required to have both a Bachelor's and a Master's degree (altogether 300 study points, referred to as credits within the European Credit Transfer System, ECTS), including periods of teaching practice within schools (Pyhältö, Pietarinen, & Salmela-Aro, 2011). Preschool teachers, whose minimum qualification requirement is a Bachelor's degree level education, are an exception.

Teacher education is one of the most popular study programmes in Finnish universities, and it typically attracts applicants with excellent secondary-level school diplomas (Malinen, Väisänen, & Savolainen, 2012). In recent years, gaining entrance to the different teacher education programmes has been an extremely competitive process. The nationwide number of applicants for primary teacher education programmes, for instance, has been approximately 7 000, and since 2011 only 10 to 11 percent of applicants have been admitted (Paronen & Lappi, 2018).

According to TALIS 2013, Finnish teachers feel that their work is valued by society, they are satisfied with their job, a vast majority of them would choose the profession again, and they trust their own expertise (OECD, 2014). Within the Finnish educational context, teachers have a high level of autonomy in their work and are trusted as professionals without being subject to inspection or accountability regimes (Sahlberg, 2011; Sarja et al., 2017).

### ***3.2. Teacher education and teaching in Germany***

Compulsory education in Germany normally starts between the age of six and seven. Depending on the federal state, children then go to primary school for four to six years, whereof the former is most common. After finishing primary school, children are distributed across three main secondary educational tracks based on their performance during their initial school years: (a) grammar school that leads to the university entry qualification (grades 5–12/13), (b) secondary modern school that leads to mid-maturity school leaving certificate (grades 5–10), and (c) general school that leads to a basic school leaving certificate (grades 5–9), entitling the holder to enter the upper secondary vocational schooling system (Döbert, 2015).

Teacher education in Germany is divided into two consecutive phases. Similar to Finland, the first phase takes place at universities where students can choose to become a teacher for either primary, general and secondary modern, or grammar schools. Each of these academic study programmes results in a Bachelor's degree and eventually a Master's or equivalent degree (altogether about 300 ECTS credits). The second, more practical phase, lasts between 18 to 24 months, and results in a state examination allowing the graduate to apply for a regular teaching position as a tenured civil servant.

Teacher education programmes are highly popular in Germany (Destatis, 2018). The popularity might be traced back to low entrance barriers at most universities for those programmes, high job security, and—in international comparison—high salaries for teachers (OECD, 2017; Weishaupt, 2014). Universities do often not require applicants to fulfil any kind of performance criteria.

Although certain studies regularly report a high level of societal respect for teachers (e.g., IfD Allensbach, 2013), the Global Teacher Status Index (Dolton, Marcenaro-Gutierrez, Pota, Boxser, & Paijpani, 2013) ranked Germany low because of low general trust in teachers' abilities to deliver good education and the common perception that students disrespect their teachers. Teaching is often associated with high job-related stress and burnout, mentioned both in the research literature and the public press (e.g., Rothland, 2013; Schaarschmidt, 2005). However, recent findings indicate that the general level of teachers' mental exhaustion—although high—is comparable with members from other social professions (Cramer, Merk, & Wesselborg, 2014). In addition, most of Germany's teachers seem to be highly satisfied with their job choice (Forsa, 2016).

#### **4. Aim of the study**

Our conceptual background suggests that motivation structures might vary across countries. The aim of our study was to explore Finnish and German undergraduate students' motivations for choosing teaching as a career. Simultaneously, we set out to validate a Finnish adaptation of the FIT-Choice scale. While this scale had already been translated into many other languages, a Finnish translation and cultural adaptation had not been undertaken prior to our own version. Accordingly, our first research question was:

RQ5. Can the FIT-Choice factor structure be replicated in a Finnish sample?

Second, it has been outlined that the situation in Finland and Germany differs extensively, as the entrance barriers in Germany are lower than in Finland and since the job situation of German teachers—that is, tenured civil servants—is more secure than for Finnish teachers. In addition, teaching is a highly respected profession in Finland, whereas the general opinion about teachers is more ambivalent in Germany. Based on this insight, we were interested in how, if at all, these differences manifest themselves in differences regarding career-choice motivations of student teachers in both countries. Thus, the second, more exploratory research question was:

RQ6. Do Finnish and German students in teacher education differ in terms of their motivation to become a teacher?

#### **5. Methods**

##### ***5.1. Sample and data collection***

The results presented here are based on a comparative study conducted with Bachelor's degree level students from two different universities in Finland and in Germany. Altogether, 413 student teachers took part in a standardised online survey. The participation was voluntary.

*The Finnish sample* ( $N_F = 217$ ) was drawn from a pool of first- to third-year students enrolled in a research-intensive university. Most of the participants studied to become subject teachers ( $n = 86$ ). The remaining sample consists of student teachers enrolled in the preschool ( $n = 57$ ) or the primary teacher programme ( $n = 51$ ). The remaining students could not be assigned to one of the three subsamples based on their answers. To reach Finnish respondents, two main strategies were used. First, several university lecturers teaching a course to the target group were asked to encourage their students to fill in the questionnaire. Second, representatives of the student organisations of the respective programmes were asked to spread the information

of the questionnaire to the target group. Due to the sampling strategy, it is unknown how many students were invited to participate in this study. For this reason, no response rate can be reported.

*The German sample* ( $N_G = 196$ ) comprised of student teachers that took part in an lecture on general educational science given at a large research-intensive university. Based on the syllabus, this lecture is usually attended by students in the second year of their degree. However, it is not uncommon that also students in their first or third year attend this lecture. The participants mostly studied to become a teacher for the grammar school track. All enrolled students were asked to fill in the questionnaire (Response rate = 75.7%).

To protect the privacy of the participants, data collection of socio-demographics was limited to gender and age. Table 2 provides an overview of the collected background information of the participants in each sample. In accordance with the gender distribution in the national teacher populations of Finland and Germany (Destatis, 2017; Statistics Finland, 2016), both samples were predominantly female.

In both samples, most participants answered all of the items in the questionnaire. However, a few students skipped large parts of the questionnaire, resulting in high numbers of missing values (*min.* = 22%, and *max.* = 93% missing values). It was decided to remove all cases with missing values to avoid model estimation problems. In the Finnish sample 14 cases and in the German sample 9 such cases were affected. This reduced the usable sample size to 390 cases in total.

**Table 2.** *Sample characteristics.*

Variable	Finnish sample	German sample
Initial $N$	217	196
Usable $N$	203	187
Female $N$	165 (81.3%)	117 (62.6%)
Age	$M = 22.76$ ( $SD = 3.47$ )	$M = 21.41$ ( $SD = 2.60$ )

## 5.2. Instrument

The FIT-Choice scale was translated into Finnish from the English original. To make sure the items maintained their original meaning and to ensure cross-language comparability, it was first translated by a certified translator. Then, the instrument was piloted in focus groups and with teacher educators who gave feedback of the functionality of the instrument. The pilot

phase yielded first evidence for both face-value validity and the applicability of the translated scale in the Finnish context. Finally, the translations were approved and verified by a bilingual Finnish team. The data of the German students were collected using the already translated and validated scale by Watt and colleagues (2012).

### 5.2.1. *Motivations for teaching*

In line with the original scale (Watt & Richardson, 2007), each FIT-Choice motivation factor was operationalised using two to five different items. Each item could be answered on a 7-point Likert scale ranging from 1 = “not at all important” to 7 = “extremely important”. All motivation items were listed in a single block introduced by the phrase: “I chose to become a teacher because...”. In addition, the block was captioned with the question: “How important were the following statements in your decision to become a teacher?”. The original items written in English, and their Finnish translations measuring the motivation factors can be found in Appendix A (the German items can be found in Watt et al., 2012).

Based on the findings by Watt and colleagues (2012), data on the motivation factor *job transferability* were not collected in the present study. First, within their study, the *job transferability* subscale exhibited insufficient psychometric properties both for their German and Norwegian (another Scandinavian country besides Finland) samples. Second, it is to be remembered that because of distinct job-entry requirements for teachers on the state level in Germany, teachers cannot easily move around within the country and, in fact, move much less often than their counterparts in other countries (Weishaupt, 2014). In other words, the subscale does not make much sense in the German or Finnish context (for more details on this argument, see Watt et al., 2012, and König & Rothland, 2012; see Kılınc et al., 2012 for a discussion on the suitability of the *job transferability* subscale in the Turkish context).

### 5.2.2. *Perceptions of teaching*

The perception factors of the FIT-Choice scale were operationalised using a set of 17 items based on a study by Watt and Richardson (2007). The items of the *task demand* and *task return* subscales were introduced by the phrase, “For each question below, please rate the extent to which you agree it is true about teaching” and followed by the question, “Do you think that...?”. The *social dissuasion* and *satisfaction with choice* items, however, were introduced by the phrase, “For each question below, please rate the extent to which it is true for you”. The respondents could answer each statement on a 7-point Likert scale ranging from 1 = “not at all” to 7 = “extremely”. Appendix B lists both the original English items and the translated Finnish equivalents (the German items can be found in Watt et al., 2012).

### 5.3. Data analyses

In line with prior cross-cultural studies using the FIT-Choice framework (Watt et al., 2012) and methodological recommendations in the literature (Davidov, Schmidt, & Billiet, 2011; Vandenberg & Lance, 2000), a multi-group confirmatory factor analysis approach was used to analyse the data. This approach enables, at the same time, testing of the psychometric quality of the measurement instrument across different samples and comparing sample differences based on latent factor scores. The former is necessary to obtain information on (a) whether a measurement instrument is suited to be used within certain cultural contexts, and (b) whether the results of those different contexts are statistically comparable. Without such information, a cross-sample comparison can easily lead to biased results (Steinmetz, 2011). The latter is required to obtain information about existing differences between samples that are not affected by measurement errors (Sass, 2011).

First, separate confirmatory factor analyses (CFA) were estimated to test the measurement model suggested by the FIT-Choice scale per sample. This test provides information on the suitability of the measurement instrument for each cultural context separately. Both CFAs were estimated using robust ML estimation with Satorra-Bentler  $\chi^2$  correction (Satorra & Bentler, 1994). To evaluate the models, the following indices and corresponding cut-off values were used (Hair, Black, Babin, & Anderson, 2014; Tabachnick & Fidell, 2007):  $CFI > .90$ ,  $TLI > .90$ ,  $RMSEA < .07$ , and  $SRMR < .08$ . In addition, coefficient  $\omega$  was used to test for congeneric reliability (McDonald, 1999). Similar to Cronbach's  $\alpha$ ,  $\omega$  should be above .60 or better .70 (Hair, Hult, Ringle, & Sarstedt, 2014).

Second, it was tested whether measurement invariance across groups can be assumed. Measurement invariance ensures that constructs (i.e., FIT-Choice factors) are measured similarly across groups and that (latent) scores of constructs are actually comparable (Meredith, 1993; Van de Vijver & Tanzer, 1997). Different forms of measurement invariance exist: (a) configural invariance, (b) metric invariance, (c) scalar invariance, and (d) measurement error invariance. Latent factor scores are only allowed to be compared between groups, as long as scalar invariance can be established. To compare the scores of manifest factors, measurement error invariance has to be fulfilled. Since the goal of this study was to compare latent factor scores, it was necessary to test whether scalar invariance would be tenable.

Testing for measurement invariance was done sequentially, whereby each model is based on the previous one (Sass, 2011). In the first step, a combined, structurally-equivalent multi-

group model was estimated without any parameter constraints (Model 1: configural invariance). In the second step, within this model, all factor loadings were constrained equally across the groups (Model 2: metric invariance). In the third step, all factor loadings and intercepts were constrained equally across the groups (Model 3: scalar invariance).  $\chi^2$  difference tests can be used to investigate whether a sequentially more restrictive measurement invariance model fits significantly worse than its nested, less restrictive predecessor. At the same time, however, it is well known that these tests are excessively sensitive in the context of complex models and larger samples (Little, 1997; Milfont & Fischer, 2010; Oishi, 2007). That is why it is often recommended (e.g., Milfont & Fischer, 2010) to rather assess what form of measurement invariance holds by assessing the fit indices described above and the changes in these indices from one model to the next. Following the arguments of Little (1997), which have also been considered in the FIT-Choice study by Watt and colleagues (2012), we decided to especially focus on the change in *TLI* during the assessment. As long as  $\Delta TLI$  is not below -.05 and all other fit indices are still above or below their corresponding cut-off values, the particular form of measurement invariance is tenable.

Third, the Finnish sample consists of student teachers within three different study programmes (preschool teacher, primary school teacher, and subject teacher). Univariate ANOVAs were used to test whether Finnish student teachers could be assumed to stem from a single population that can be compared to the German sample. This was expected since prior research shows that student teachers within different study tracks do not differ in their personality (Faust & Förster, 2008) nor in their socio-economic backgrounds (Kühne, 2006). However, it was necessary to test for sample homogeneity, as otherwise the country comparison might have shown non-significant results for certain scales that were masked because of relevant differences within the Finnish sample.

Fourth, to answer the research question of whether Finnish and German students in teacher education differ in terms of their motivation to become a teacher, a latent mean analysis (LMA) was conducted (Brown, 2015). The scalar invariance model (Model 3), in which the latent mean of the German group was set equal to zero, was used for this purpose. This allowed for freely estimating the latent mean of the Finnish group and to test for differences. The obtained estimate represents the latent factor mean difference ( $\Delta M$ ) between both groups, with the German sample as a reference. The German sample was chosen as the reference group since we are interested in how Finnish students differ from their German counterparts. Scaling was done using the marker variable strategy selecting the item as a marker that exhibited strongest invariance between both groups (Brown, 2015; McCoach, Gable, &

Madura, 2013). For each FIT-Choice factor, several configural invariance models with different items as marker variables were estimated for this purpose (Brown, 2015; McCoach, Gable, & Madura, 2013). This iterative procedure allowed us to choose the most invariant item required for scaling the metric of each latent factor. Cohen's  $d$  is reported as effect size measure transferring  $\Delta M$  into a common and familiar metric. In line with Glass (1976) and Keselman, Algina, Lix, Wilcox, and Deering (2008), the standard deviation of the reference group ( $SD_{German}$ ) was used to standardise the mean difference in  $d$ . The LMA results will only be reported for scales that have been answered homogeneously within the Finnish sample or where the descriptive mean structure of the three different Finnish subsamples prove to be either consistently lower or consistently higher than the total mean of the German sample. This way it is made sure that the overall means of both countries are indeed comparable and differences are not masked because of overlapping mean structures.

All analyses were conducted using *R* 3.4.3 (R Core Team, 2017), including the packages *lavaan* 0.5–23 (Rosseel, 2012), *semTools* 0.5–1 (Jorgensen, Pornprasertmanit, Schoemann, & Rosseel, 2018), *psych* 1.7–5 (Revelle, 2017), and *agricolae* 1.2–8 (de Mendiburu, 2017).

## 6. Results

### 6.1. Initial factor structure of motivation and perception factors for each country

Our first aim was to investigate whether the FIT-Choice factor structure can be replicated with a Finnish sample (RQ1). In the first step, separate CFAs for the *motivations* and *perceptions* factors for both countries were estimated. Those CFAs enabled identifying items that exhibited psychometric issues in either of the samples. First, factor loadings below .50 were observed for items B7 (*intrinsic value*; both samples), B10 (*work with children/adolescents*; German sample), and B31 (*make social contribution*; Finnish sample). Consequently, all three items were removed from any further analysis. Second, item D3 (*satisfaction*) caused a negative variance in the Finnish sample. Because the *satisfaction* factor was only measured by two items (D3, D5), the whole scale had to be removed from any further analysis. In the second step, all four CFAs were estimated again without the problematic items and showed an acceptable-to-good fit (Table 3).

**Table 3.** CFA results without problematic items.

Model	$\chi^2_{SB}$	df	RMSEA	SRMR	CFI	TLI
<i>Finnish sample</i>						
Motivations	540.98***	379	.046	.060	.929	.913
Perceptions	118.56**	80	.049	.053	.960	.948
<i>German sample</i>						
Motivations	532.30***	379	.047	.066	.931	.915
Perceptions	109.94*	80	.045	.054	.960	.948

Note. \*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$ . FIT-Choice framework without the following items: B7, B10, B31, D3, and D5.

Table 4 lists the composite reliabilities for all subscales in both samples. Similar to Cronbach's  $\alpha$ ,  $\omega$  were expected to be above .70 or at least .60. These cut-off values were met by most of the scales. However, *shape future of children/adolescents*, *make social contribution*, *fallback career*, *expert career*, and *social dissuasion* exhibited relatively low  $\omega$  in at least one sample, usually accompanied by a sufficiently high  $\omega$  in the respective other sample. It was decided to not exclude any scale, since the reliabilities could be judged as sufficient. In addition, the CFAs exhibited a good fit of the measurement model. For further reference, Appendix C and D present the correlations between the *motivations*, as well as the *perceptions* factors.

To sum up, based on the results of the CFA and the composite reliabilities, the factor structure of the FIT-Choice framework could be both replicated in the Finnish and the German sample. However, a few single items of the original instrument had to be omitted from any further analysis due to psychometric problems.

**Table 4.** Composite reliabilities.

Factors	$\omega$	
	Finnish sample	German sample
Motivations		
<i>Perceived teaching abilities</i>	.83	.80
<i>Intrinsic value</i>	.79	.82
<i>Job security</i>	.75	.86
<i>Time for family</i>	.83	.82
<i>Shape future of children/adolescents</i>	.68	.63
<i>Enhance social equity</i>	.75	.71

<i>Make social contribution</i>	.60	.72
<i>Work with children/adolescents</i>	.94	.92
<i>Fallback career</i>	.62	.69
<i>Prior teaching and learning experiences</i>	.88	.84
<i>Social influences</i>	.85	.87
<hr/>		
Perceptions		
<i>High demand</i>	.78	.71
<i>Expert career</i>	.61	.79
<i>Social status</i>	.84	.84
<i>Salary</i>	.94	.89
<i>Social dissuasion</i>	.68	.59

Note.  $\omega$  = Composite reliability coefficient based on McDonald (1999).

## 6.2. Validation of motivation and perception factors across countries

Next, we examined measurement invariance across both samples with the FIT-Choice scale to compare the latent means of these factors between Finland and Germany (RQ2). A summary of the test for measurement invariance between both countries, incorporating the preliminary modifications, can be found in Table 5. The  $\chi^2$  difference tests indicate significant differences between almost all of the tested measurement models but between Model 1b and 2b. As foreshadowed, however,  $\chi^2$  tests are overly sensitive and therefore give only limited information on whether measurement invariance can be assumed. We therefore decided to rather focus on fit indices and their changes between the sequentially constrained models to assess measurement invariance issues. Based on these criteria, both configural and metric invariance can be established for the *motivations* and *perceptions* subscales (changes from Model 1a/1b to 2a/2b). However, the changes in the fit indices from Model 2a/2b to 3a/3b indicate that full scalar invariance cannot be assumed for either the *motivations* or *perceptions* subscales, since  $\Delta TLLM = -.072$  and  $\Delta TLLP = -.119$  lie far below the considered cut-off value of  $-.05$ . This suggests that one or more of the items functioned differently between the samples (i.e., the response styles between the countries may have been different). More detailed investigations revealed that the constrained intercepts of the items B6 and B20 (*make social contribution*), B9 and B23 (*shape future of children/adolescents*), B18 and B29 (*time for family*), B24 (*social influences*), B27 (*job security*), C8 (*social status*), as well as C14 and C10 (*expert career*) caused the low fit values. In fact, after freeing the equality constraints of the intercepts of these eleven items, partial scalar invariance could be

established for both the *motivations* and *perceptions* subscales (changes from Model 2a/2b to 4a/4b resulted in  $\Delta TLI > -.05$ ).

Partial scalar invariance is a sufficient condition for the comparison of factor means, as long as the number of non-invariant items is small in comparison to the remaining invariant items operationalising a latent construct (Sass, 2011; also Bryne, Shavelson, & Muthén, 1989; Steenkamp & Baumgartner, 1998). Thus, the Finnish and German means of the factors *time for family*, *social influences*, *job security*, and *social status* can be compared using LMA without substantial issues (see Appendix A and B for the number of items in each scale). However, mean differences between both samples in regard to the factors *make social contribution*, *shape future of children/adolescents*, and *expert career* have to be interpreted carefully. These three factors were more strongly affected by the identified measurement invariance problems, since more than half of the items of the respective subscales exhibited considerable differences between both samples that cannot be traced back to the underlying latent construct.

**Table 5.** Summary of fit indices for invariance tests.

Model	$\chi^2_{SB}$	df	$p_{\chi^2}$	RMSEA	SRMR	CFI	TLI	$\Delta TLI$
Motivations								
1a. Configural invariance	951.72	700		.043	.060	.942	.927	
2a. Metric invariance	982.18	719	.012	.043	.063	.939	.926	-.001
3a. Scalar invariance	1 272.08	738	<.001	.061	.072	.876	.854	-.072
4a. Partial scalar invariance <sup>1</sup>	1 104.06	730	<.001	.051	.068	.913	.897	-.029
Perceptions*								
1b. Configural invariance	230.61	161		.047	.054	.959	.947	
2b. Metric invariance	238.11	171	.755	.045	.055	.961	.952	.005
3b. Scalar invariance	426.62	181	<.001	.083	.073	.856	.833	-.119
4b. Partial scalar invariance <sup>2</sup>	288.64	178	<.001	.056	.061	.935	.923	-.029

Note.  $p_{\chi^2}$  =  $p$ -value for  $\chi^2_{SB}$  difference test between nested models (i.e., Models 2 vs 1, 3 vs 2, 4 vs 2) based on the method proposed by Satorra and Bentler (2010). \* The factor loadings of both items of the expert career factors had to be constrained to be equal.<sup>1</sup> Intercepts of the following items were allowed to vary between both groups: B6, B9, B18, B20, B23, B24, B27, and B29. <sup>2</sup> Intercepts of the following items were allowed to vary between both groups: C8, C10, and C14.

### 6.3. Different motivations for and perceptions of teaching across countries

To answer our second research question, latent mean differences between both countries regarding *motivations* for teaching and *perceptions* of teaching were compared. However, the Finnish sample consists of three different study tracks. In the first step, it was therefore necessary to check whether the Finnish students could be assumed to stem from a

homogeneous sample. For this purpose, ANOVAs were calculated to test whether the three Finnish subsamples (preschool, primary school, and subject teachers) differed significantly from each other. The results of these tests including effect sizes ( $\eta^2$ ), the descriptive means ( $M$ ) and standard deviations ( $SD$ ) for all three subsamples, the Finnish overall mean ( $M_F$ ) and the German overall mean ( $M_G$ ) can be found in Table 6. In the second step, a LMA was estimated to test whether Finnish and German students systematically differed in their *motivations* for teaching and *perceptions* of teaching for all scales to be found homogeneous. Significant differences between the Finnish subsamples could be found for two of the eleven motivation scales (*time for family*,  $p < .001$ , medium effect; *prior teaching and learning experiences*,  $p = .049$ , small effect) and three of the five perceptions scales (*expert career*,  $p = .011$ , small effect; *social status*,  $p < .001$ , medium effect; *salary*,  $p < .001$ , medium effect). This suggests that for these scales it cannot be assumed that the three Finnish subsamples stem from a homogeneous sample. However, a closer look reveals that the means of these subsamples are consistently lower than the mean of the German subsample for *time for family* and *salary*. In other words, despite the differences within the Finnish sample, both scales can still be compared using the LMA framework because the mean structure follows a clear tendency for these two scales and there is no risk that differences between both countries are masked because of sample heterogeneity reasons. The remaining three scales, *prior teaching and learning experiences*, *expert career*, and *social status*, will not be used in the LMA framework due to their heterogeneous mean structure in the Finnish sample that is not consistently lower or higher than the German overall mean.

Post-hoc Scheffé tests were used to test how students within the different Finnish study tracks differed from each other for those scales that showed a significant ANOVA test. Concerning the scales *time for family* and *social status*, differences could be found between preschool student teachers and both subject and primary school student teachers ( $p < .05$ ). However, no differences could be found between subject and primary school student teachers for these two scales. In regard to the scale *prior teaching and learning experience*, differences could be found between subject student teachers and preschool student teachers ( $p < .05$ ) but not between subject and primary student teachers as well as preschool and primary student teachers. The observed heterogeneity in the scale *expert career* can mainly be traced back to differences between preschool and the primary student teachers ( $p < .05$ ). Both groups, however, showed no differences to subject teachers. Concerning the scale *salary*, differences could be found between all three groups ( $p < .05$ ).



**Table 6.** ANOVA comparison between Finnish subsamples including overall descriptives for the Finnish and German sample.

	Finnish sample					German sample	
	Preschool	Primary	Subject	ANOVA		Total	Total
	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>	<i>p</i>	$\eta^2$	<i>M<sub>F</sub> (SD)</i>	<i>M<sub>G</sub> (SD)</i>
<b>Motivations</b>							
<i>Perceived teaching abilities</i>	5.42 (0.98)	5.68 (0.86)	5.45 (0.96)	.163	0.010	5.51 (0.94)	5.76 (0.77)
<i>Intrinsic value</i>	6.11 (0.93)	6.27 (0.77)	6.00 (0.94)	.356	0.004	6.11 (0.89)	6.07 (0.86)
<i>Job security</i>	4.60 (1.30)	4.56 (1.25)	4.71 (1.26)	.889	0.000	4.65 (1.24)	5.34 (1.14)
<i>Time for family</i>	2.30 (1.11)	3.28 (1.58)	3.47 (1.41)	<.001	0.067	3.06 (1.44)	4.03 (1.27)
<i>Shape future of children/ adolescents</i>	5.75 (1.15)	5.53 (1.03)	5.13 (1.39)	.303	0.006	5.43 (1.25)	5.85 (0.90)
<i>Enhance social equity</i>	5.81 (1.22)	5.68 (1.34)	5.26 (1.26)	.535	0.002	5.56 (1.28)	5.57 (1.08)
<i>Make social contribution</i>	5.10 (1.26)	5.08 (1.38)	4.77 (1.35)	.896	0.000	4.93 (1.34)	5.62 (1.04)
<i>Work with children/ adolescents</i>	6.49 (0.80)	6.29 (1.01)	4.92 (1.51)	.323	0.005	5.77 (1.39)	5.77 (1.11)
<i>Fallback career</i>	1.62 (0.91)	1.59 (0.85)	1.86 (1.06)	.946	0.000	1.73 (0.98)	2.12 (1.17)
<i>Prior teaching and learning experiences</i>	4.81 (1.40)	5.29 (1.52)	5.72 (1.08)	.049	0.020	5.32 (1.36)	4.92 (1.38)
<i>Social influences</i>	3.43 (1.75)	3.54 (1.28)	3.41 (1.73)	.721	0.001	3.43 (1.62)	4.42 (1.50)
<b>Perceptions</b>							
<i>High demand</i>	5.92 (0.75)	5.84 (0.87)	5.55 (0.85)	.573	0.002	5.74 (0.83)	5.75 (0.80)
<i>Expert career</i>	5.70 (0.81)	5.29 (0.90)	5.45 (0.79)	.011	0.034	5.48 (0.83)	5.60 (1.06)
<i>Social status</i>	4.00 (0.86)	4.83 (0.88)	4.63 (0.86)	<.001	0.114	4.52 (0.92)	4.60 (1.00)
<i>Salary</i>	2.47 (1.08)	3.67 (1.10)	4.22 (1.20)	<.001	0.119	3.57 (1.33)	5.09 (0.93)
<i>Social dissuasion</i>	2.68 (1.04)	2.64 (1.12)	2.76 (1.15)	.877	0.000	2.70 (1.10)	3.59 (1.20)

The results of the LMA can be found in Table 7.<sup>1</sup> The  $\Delta M$  values describe the mean difference between both countries regarding the particular FIT-Choice factors.

Most of the observed significant differences between the two countries were negative. On average, Finnish students judged *perceived teaching abilities* ( $p = .006$ , small effect), *job security* ( $p < .001$ , medium effect), *time for family* ( $p < .001$ , large effect), *fallback career* ( $p < .001$ , medium effect), *social influences* ( $p < .001$ , medium effect), *salary* ( $p < .001$ , large effect), and *social dissuasion* ( $p < .001$ , large effect) as less important for their decision to become a teacher than their German counterparts did. Significant differences between both samples, however, could not be found for the remaining FIT-Choice factors.

**Table 7.** Results of latent mean analysis.

Factors	$\Delta M$	$z$	$p$	$d$
Motivations				
<i>Perceived teaching abilities</i>	-0.23	-2.745	.006	-0.34
<i>Intrinsic value</i>	0.02	0.29	.772	0.03
<i>Job security</i>	-0.54	-4.22	.000	-0.51
<i>Time for family</i>	-1.25	-5.68	.000	-0.95
<i>Shape future of children/adolescents</i>	0.00	0.01	.989	0.00
<i>Enhance social equity</i>	0.01	0.10	.923	0.01
<i>Make social contribution</i>	0.00	-0.05	.962	0.00
<i>Work with children/adolescents</i>	0.00	-0.03	.975	0.00
<i>Fallback career</i>	-0.49	-4.519	.000	-0.49
<i>Prior teaching and learning experiences</i>	NA	NA	NA	NA
<i>Social influences</i>	-0.82	-4.74	.000	-0.54
Perceptions				
<i>High demand</i>	-0.07	-0.76	.448	-0.09
<i>Expert career</i>	NA	NA	NA	NA
<i>Social status</i>	NA	NA	NA	NA
<i>Salary</i>	-1.50	-12.82	.000	-1.74
<i>Social dissuasion</i>	-0.66	-5.73	.000	-0.95

*Note.* The German sample was used as the reference group. A positive (negative)  $\Delta M$  therefore indicates that Finnish students have higher (lower) latent means regarding the FIT-Choice factors in comparison to their German counterparts.  $z$  = Wald test statistic. NA = Not available because the Finnish sample is not homogeneous or its mean structure is not consistently lower or higher than the German overall mean (Table 6).

1 Deviations between the depicted mean differences in Table 6 and Table 7 are due to the fact that the observed factor means are still affected by measurement errors while the latent factor means are not.

## 7. Discussion and conclusions

Finland and Germany have different education systems, teacher education programmes and study entrance criteria, as well as different perceptions of teaching as a profession. One aim of this study was to explore whether the outlined differences between both countries manifest themselves in students' motivations for becoming a teacher and in their perceptions of this particular career choice. For this purpose, the FIT-Choice scale was translated into Finnish to explore Finnish student teachers' career-choice motivations, and to compare the motivations for becoming a teacher between the Finnish and German students who were enrolled in teacher education programmes. Accordingly, the validation of the Finnish scale is discussed next (Section 7.1), followed by the exploration of differences that could be found between the three different study tracks within the Finnish sample (Section 7.2), and the differences between the Finnish and the German samples (Section 7.3). After that, a discussion regarding the limitations and implications of this study will be provided (Section 7.4).

### 7.1. A replicable FIT-Choice factor structure for Finland

This study explored whether the theoretically developed and manifoldly validated FIT-Choice structure can also be replicated with a Finnish sample. Overall, the analyses suggest that the FIT-Choice framework is applicable to the Finnish context. Although the results are generally consistent with the FIT-Choice framework as presented and validated by other researchers in other cultural contexts (e.g., Fokkens-Bruinsma & Canrinus, 2012; Watt et al., 2012), a few minor differences were nonetheless found. First, two items, namely B7 (*intrinsic career value*; "I have always wanted to be a teacher") and B31 (*make social contribution*; "Teaching enables me to give back to society") exhibited substantially low factor loadings, indicating that they do not reflect the underlying latent constructs well. Second, the factor *satisfaction with choice* could not be replicated, since item D3 ("How satisfied are you with your choice of becoming a teacher?") caused serious estimation problems in the CFA approach (Heywood case; Kline, 2016).

Interestingly, item B7 also exhibited certain psychometric problems in other studies. Watt and colleagues (2012) had to omit the item to improve the reliability of its corresponding subscale. Although the authors did not describe what caused the low subscale reliability, a low observed factor loading could be a potential reason. Relatively low factor loadings of item B7 have also been reported in a Chinese sample by Lin and colleagues (2012), as well as in a Dutch sample by Fokkens-Bruinsma and Canrinus (2012). Furthermore, item B31 caused issues in the study by Kılınç and colleagues (2012). To be more concrete, the item had to be

deleted because of factor cross-loadings. Taken together, these findings suggest general psychometric problems with respect to items B7 and B31. The estimation problem caused by item D3, however, points more to cultural and professional particularities. Problems with this item have not yet been documented in the literature.

## **7.2. Differences within the Finnish sample**

The Finnish sample consists of students studying to become subject teachers as well as students that aim to become preschool and primary school teachers. No differences between these three subsamples could be found for most scales in the FIT-Choice framework, but *time for family, prior teaching and learning experiences, expert career, social status, and salary*. Further analyses showed that it is the group of preschool teachers that most strongly differs from both the primary school and the subject teachers. Preschool student teachers indicated less often that they had chosen teaching because they thought the job offers much time for family duties and because of prior teaching and learning experiences in comparison to both other groups. In addition, they tend to less strongly perceive teaching as a job that is socially valued and respected or as a job that earns a comparatively good remuneration.

These findings are not completely surprising. First of all, students visit preschool at a very young age and many aspiring teachers might not remember their own time within this part of the education system. That could be the reason why preschool student teachers less often stress positive role models and positive learning experiences as motivations for their career choice in comparison to primary school and subject teachers. Second, preschool teachers' average salary in Finland is in fact considerably lower than the ones of subject teachers and primary teachers (Statistics Finland, 2018). It follows that salary should be less of an incentive to become a teacher for the preschool system. Third, previous studies (Eskelinen & Hjelt 2017; Perho & Korhonen 2012) have shown that preschool teachers perceive a lack of appreciation of their profession and challenging work environment. This might explain why the preschool student teachers perceive their future job as less socially valued and respected than primary and subject student teachers.

## **7.3. Cross-cultural comparison**

### **7.3.1. Relevance of motivations and perceptions**

The sample comparisons showed some similarities between both countries in terms of the relevance of the motivations and perceptions. First, both Finnish and German students rated *intrinsic value* highest as their motivation to become teachers. Second, *perceived teaching abilities, enhance social equity, work with children/adolescents, and shape the future of*

*children/adolescents* were also consistently rated as important motivations behind the participants' job choices in both samples. Third, most of the students in both samples indicated that they chose teaching as a first-choice alternative and not as a *fallback career*. Fourth, teaching was perceived by both Finnish and German students as a job with *high demands* that requires a range of professional *expert knowledge*. Fifth, most of the students in both samples did not experience much *social dissuasion* when it came to their career choice. Not only were Finnish and German students similar with respect to the just-described motivations and perceptions, but these findings also strongly replicate those of other studies with very similar results (e.g., Fokkens-Bruinsma & Canrinus, 2012; Hennessy & Lynch, 2017; König & Rothland, 2012; Watt et al., 2012; Watt & Richardson, 2007). In other words, a lot of evidence exists that certain motivations for becoming a teacher and particular perceptions regarding the teaching career are quite universal in nature (at least in Western culture; for slightly different findings in an Asian sample, see, Lin et al., 2012).

### 7.3.2. *Observed differences*

Although some similarities could be found between both countries in terms of students' motivations for and perceptions of teaching, several differences can also be observed between the samples. In comparison to the Finnish students, German students tended to choose their career in more cases because they (a) perceived their *teaching abilities* as high, (b) judged the job as having a high *personal utility value* including strong *job security* and *time for family*, and (c) were convinced by other people that this career choice would be a good idea.

Although most students in both samples indicated that they did not choose teaching as a *fallback career*, the German students scored significantly higher on this factor than their Finnish counterparts did. Finnish students, however, reported less *dissuasion* from their social environment regarding their choice to pursue teaching as a career, indicating a higher rate of social acceptance concerning the choice to become a teacher. Another meaningful difference could be observed in terms of the perception of teacher remuneration. Finnish students did not judge teaching to be as well paid as the German students thought they are.

### 7.3.3. *Explaining the differences between Finland and Germany*

To sum up, this study found that German student teachers base their career choices more strongly on extrinsic motives (e.g., *job security*, *expected future salary*, *time for family*), or on other more external reasons (e.g., *social influences* or a lack of other options and thus as a *fallback solution*) than Finnish students do. In addition, this study also found that German

student teachers are more often socially discouraged to take up teaching as a career although they judge the future earning opportunities to be quite good.

Some of these findings might be traced back to the different (teacher) education systems in both countries. In Finland, teaching is a highly respected profession and rigorous admission criteria make sure that only the best students are allowed to enter the teacher education programme. In Germany, however, teaching is a less respected profession and entrance barriers to become a teacher are much lower. At the same time, German teachers are usually tenured civil servants with long-term job security from the outset, while Finnish teachers often start with fixed-term contracts that only later develop into tenured positions with high job security either as municipal civil servants or employees. Teachers' average salary also differs between the two countries. For instance, while an experienced German upper secondary education teacher earns about 68 000 EUR annually, her or his Finnish counterpart earns around 42 000 EUR (OECD, 2017).

These structural differences might explain why personal utility values found greater appreciation in the German sample, whereas most of the social utility values were judged to be quite similar in both samples. On the one hand, the Finnish system might attract mostly students who are motivated to fulfil their intrinsic social utility values. The Finnish system, however, does not attract students who only try to maximise personal utility values. Especially the high entrance barriers in combination with the comparably lower material incentives, such as salary, might actually discourage students that only look for a job that offers attractive benefits. The German system, on the other hand, might attract a wider range of students. Besides the strongly intrinsically motivated students, other candidates might choose to study teacher education to get a job that offers high job security, time for family, and is well paid. Some of these students might even choose to become a teacher only because of a failure to pursue another, first-choice career. After all, strong entrance barriers to teacher education do not exist in Germany.

In regard to the interpretations about the status of the teaching profession in the Finnish and German societies, the findings are difficult to judge. As described earlier, teaching is a highly respected profession in Finland, whereas the situation is more ambivalent in Germany. Our findings do not fully reflect this presupposition. On the one hand, both Finnish and German teacher education students similarly indicated that they perceive teachers as having a moderate level of social status and respect. On the other hand, German students experienced a higher level of *social dissuasion* than Finnish students. However, the factor means were still

below the middle of the scale. Hence, the students' judgement does not indicate a high level of societal rejection.

#### ***7.4. Limitations and implications of this study***

This study has certain limitations. Most importantly, all analyses are based on two convenience samples of relatively small size stemming from only one single Finnish and one German university. First, it cannot be fully ruled out that those two universities are to some extent idiosyncratic and that the sample therefore does not necessarily represent students from other universities in the respective countries. Second, to protect the privacy of students within the sample, no other socio-demographic data than age and gender were collected. It follows that this study was not able to provide information on how student teachers' individual backgrounds (e.g., enrolment year, socio-economic status) were connected to their motivations for teaching. It cannot be ruled out that some of these variables confound the findings of the cross-cultural comparison reported here. Third, although no detailed information about the individual study participants is available, it is known that the German sample consists of students studying exclusively for the Grammar school track and the Finnish sample includes students from preschool, primary school, and subject teacher programmes.

The ANOVA tests showed that for most FIT-Choice scales the Finnish sample could indeed shown to be homogeneous. Nevertheless, the students of the three Finnish study tracks differ to a certain extent in their career choice motivations and in their perceptions about teaching as a career. Because of the small size of the Finnish sample, these differences including their effect on the psychometric characteristics of the Finnish FIT-Choice scales could not be analysed using elaborate statistical techniques. Moreover, the German sample did not contain any students outside the Grammar school track. Thus it was not possible to conduct a more detailed comparison between students from these two countries within comparable teacher education programmes (e.g., primary school student teachers).

Based on these limitations, future studies should aim to replicate our findings using samples that (a) are larger, (b) stem from more than one university in each country, (c) include a sufficient number of students of each teacher education programme, and (d) collect more personal information that can be used as control variables or that are subject of further separate analyses. Especially the latter might enable the disentanglement of cultural and individual effects on students' motivation to become teachers, including their perceptions about a career in teaching that might be masked by our findings.

Another important concern regarding this study refers to social desirability. Social desirability describes the phenomenon that study participants answer items to present themselves in a favourable manner relative to prevailing social norms and thereby threaten the validity of empirical findings (King & Bruner, 2000). In this study, it is the factor *fallback career* in particular that is probably prone to such a bias, as well as, to a more moderate extent, the factors *intrinsic career value* and *time for family*. However, it cannot be determined whether our findings are indeed affected by this kind of bias and to what extent. Additionally, it remains open whether both samples were similarly affected by such potential issues. Future studies using the FIT-Choice framework should therefore include existing items that allow to detect social desirability issues within the data (e.g., Stöber, 2001).

Although this study contributed to the list of comparisons of motivations for becoming a teacher between several countries, the interpretation of this cross-cultural comparison should be considered with care, since measurement invariance issues could not be ruled out entirely. Within the literature, different authors proposed different criteria to assess whether scales are invariant between two samples. In this study, measurement invariance was assessed via absolute fit indices (*RMSEA*, *SRMR*, *CFI*, *TLI*), and changes of *TLI* from one tested model to another. Other criteria including overly sensitive  $\chi^2$  difference tests that indicated significant differences between the different measurement models were reported but disregarded. We hereby mainly follow the suggestions which have also been used in the comparative FIT-Choice study by Watt et al. (2012). Although we recognise that other scholars would have focussed on other criteria and therefore could have come to other interpretations concerning the invariance of the used instrument in this study, the chosen approach allows our findings to be compared to the findings of Watt and colleagues. Moreover, based on the chosen approach, full scalar invariance could not be established for a few factors of the FIT-Choice scale. Full scalar invariance is, however, required to compare latent means between two samples. Therefore, without full scalar invariance it remains open whether the same latent construct is measured across the samples of interest. Especially the factors *make social contribution*, *shape future of children/adolescents*, and *expert career* were affected by this issue. Future studies should further focus on measurement invariance issues of the FIT-Choice framework in the context of cross-cultural studies.

Before this study, no Finnish adaptation of the FIT-Choice framework existed. It follows that this study contributed to the literature by developing and validating such an adaptation. All in all, the findings of this study imply that the newly developed version of the FIT-Choice scale is suitable for research in the Finnish context. However, it is unclear whether the

psychometric problems, as identified in this study, can be traced back to the Finnish cultural context as assumed here. It cannot be ruled out that the psychometric issues are sample-specific since all Finnish data in this study stems from only one university. It follows that future studies should not omit the affected items but include them in their survey instruments. This would then also enable researchers to see whether the findings presented here can be replicated. In addition, the Finnish sample only contained about 200 students. That is why all analyses have been conducted with the whole sample and not the three different subsamples of students studying within different teacher programmes. It would be interesting for future studies to test whether the Finnish version of the FIT-Choice instrument works equally well for all teacher tracks in Finland. Some of the identified psychometric problems might also be traced back to different answer behaviour of students in these different study programmes.

It is an important endeavour to investigate student teachers' career choice motivations, as has been done in this study. However, it might be even more important to understand how such motivations affect certain outcome criteria, such as study performance, later teaching performance, job satisfaction, and psychological well-being in the long run. Future research on the impacts on such variables would be of great relevance.

### **Funding**

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

### **Competing interests statement**

The authors have no competing interests to declare.

### **References**

- Akar, E. O. (2012). Motivations of Turkish pre-service teachers to choose teaching as a career. *Australian Journal of Teacher Education, 37*(10), 67–84.
- Brown, T. A. (2015). *Confirmatory factor analysis for applied research*. New York: Guilford Press.
- Byrne, B. M., Shavelson, R. J., & Muthén, B. (1989). Testing for equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin, 105*, 456–466.
- Cramer, C., Merk, S., & Wesselborg, B. (2014). Psychische Erschöpfung von Lehrerinnen und Lehrern: Repräsentativer Berufsgruppenvergleich unter Kontrolle berufsspezifischer Merkmale [Mental exhaustion of teachers: Representative comparison of occupational groups controlled for profession-specific variables]. *Lehrerbildung auf dem Prüfstand, 7*(2), 138–156.
- Davidov, E., Schmidt, P., & Billiet, J. (Eds.) (2011). *Cross-cultural analysis: Methods and applications*. New York: Routledge.

- De Mendiburu (2017). *agricolae: Statistical procedures for agricultural research*. R package version 1.2-8. <https://CRAN.R-project.org/package=agricolae>
- Destatis (2017). *Bildung und Kultur: Allgemeinbildende Schulen – Schuljahr 2016/2017 (Fachserie 11, Reihe 1)* [Education and culture: General education schools – Term 2016/2017.]. Wiesbaden: Statistisches Bundesamt [The Federal Statistical Office of Germany].
- Destatis (2018). *Sonderauswertung: Studierende in Lehramtsstudiengänge* [Special analysis: Students studying teaching education]. Wiesbaden: Statistisches Bundesamt [The Federal Statistical Office of Germany]
- Dolton, P., Marcenaro-Gutierrez, O., Pota, V., Boxser, M., & Pajjani, A. (2013). Global teacher status index 2013. London: Varkey GEMS Foundation. Retrieved July 25, 2018, from <https://www.varkeyfoundation.org/media/2787/2013globalteacherstatusindex.pdf>
- Döbert, H. (2015). Germany. In W. Hörner, H. Döbert, L. R. Reuter, & B. von Kopp (Eds.), *The education systems of Europe* (pp. 305–333). Cham: Springer.
- Eccles, J. S., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L., & Midgley, C. (1983). Expectancies, values, and academic behaviors. In J. T. Spence (Ed.), *Achievement and achievement motivation* (pp. 75–146). San Francisco, CA: W. H. Freeman.
- Erss, M. (2018). “Complete freedom to choose within limits” – Teachers’ views of curricular autonomy, agency and control in Estonia, Finland and Germany. *The Curriculum Journal*, 29(2), 238–256.
- Eskelinen, M., & Hjelt, H. (2017). Varhaiskasvatuksen henkilöstö ja lapsen tuen toteuttaminen. Valtakunnallinen selvitys 2017 [Early childhood education staff and implementing support for children. A national report 2017]. Helsinki: Ministry of Education and Culture.
- Faust, G., & Foerster, F. (2008). Offenheit für Ideen und Handlungen bei Grundschullehramtsstudierenden [Openness for ideas and activities of primary-school teachers]. In A. Hartinger, R. Bauer, & R. Hitzler (Eds.), *Veränderte Kindheit: Konsequenzen für die Lehrerbildung* (pp. 129-139). Bad Heilbrunn: Klinkhardt.
- Fokkens-Bruinsma, M., & Canrinus, E. T. (2012). The factors influencing teaching (FIT)-Choice scale in a Dutch teacher education program. *Asia-Pacific Journal of Teacher Education*, 40(3), 249–269.
- Forsa (2016). Zufriedenheit im Lehrerberuf: Ergebnisse einer repräsentativen Lehrerbefragung. [Satisfaction in the teaching profession: Findings of a representative teacher survey]. Berlin: forsa Politik- und Sozialforschung GmbH.
- Glass, G. V. (1976). Primary, secondary, and meta-analysis research. *Educational Researcher*, 5, 3–8.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2014). *Multivariate data analysis* (7th ed.). Harlow: Pearson.
- Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2014). *A primer on partial least squares structural equations modeling (PLS-SEM)*. Los Angeles: SAGE.
- Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Abingdon, UK: Routledge.

- Hennessy, J., & Lynch, R. (2017). "I chose to become a teacher because". Exploring the factors influencing teaching choice amongst pre-service teachers in Ireland. *Asia-Pacific Journal of Teacher Education*, 45(2), 106–125.
- IfD Allensbach (2013). *Allensbacher Berufsprestige-Skala 2013* [Occupational Prestige scale 2013]. Allensbach: IfD Allensbach.
- Jorgensen, T. D., Pornprasertmanit, S., Schoemann, A. M., & Rosseel, Y. (2018). *semTools: Useful tools for structural equation modeling*. R package version 0.5-1. Retrieved from <https://CRAN.R-project.org/package=semTools>
- Jugovic, I., Marusic, I., Ivanec, T. P., & Vidovic, V. V. (2012). Motivation and personality of preservice teachers in Croatia. *Asia-Pacific Journal of Teacher Education*, 40(3), 271–287.
- Keselman, H. J., Algina, J., Lix, L. M., Wilcox, R. R., & Deering, K. N. (2008). A generally robust approach for testing hypotheses and setting confidence intervals for effect sizes. *Psychological Methods*, 13(2), 110–129.
- Kılınç, A., Watt, H. M. G., & Richardson, P. W. (2012). Factors influencing teaching choice in Turkey. *Asia-Pacific Journal of Teacher Education*, 40(3), 199–226.
- King, M. F., & Bruner, G. C. (2000). Social desirability bias: A neglected aspect of validity testing. *Psychology and Marketing*, 17(2), 79–103.
- Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th ed.). New York: The Guilford Press.
- König, J., & Rothland, M. (2012). Motivations for choosing teaching as a career: Effects on general pedagogical knowledge during initial teacher education. *Asia-Pacific Journal of Teacher Education*, 40(2), 89–315.
- Kühne, S. (2006). Das soziale Rekrutierungsfeld der Lehrer. Empirische Befunde zur schichtspezifischen Selektivität in akademischen Berufspositionen [The social field recruiting of teachers. Empirical findings concerning social class-related selectivity in academic jobs]. *Zeitschrift für Erziehungswissenschaft*, 9, 617-631.
- Kyriacou, C., & Coulthard, M. (2000). Undergraduates' views of teaching as a career choice. *Journal of Education for Teaching: International research and pedagogy*, 26(2), 117–126.
- Lin, E., Shi, Q., Wang, J., Zhang, S., & Hui, L. (2012). Initial motivations for teaching: Comparison between preservice teachers in the United States and China. *Asia-Pacific Journal of Teacher Education*, 40(3), 227–248.
- Little, T. D. (1997). Mean and covariance structures (MACS) analyses of cross-cultural data: Practical and theoretical issues. *Multivariate Behavioural Research*, 32(1), 53–76.
- Malinen, O.-P., Väisänen, P., & Savolainen, H. (2012). Teacher education in Finland: A review of a national effort for preparing teachers for the future. *The Curriculum Journal*, 23(4), 567–584.
- Manuel, J., & Hughes, J. (2006). "It has always been my dream": Exploring preservice teachers' motivations for choosing to teach. *Teacher Development*, 10(1), 5–24.

- McCoach, D. B., Gable, R. K., & Madura, J. P. (2013). *Instrument development in the affective domain: School and corporate applications* (3rd ed). New York: Springer.
- McDonald, R. P. (1999). *Test theory: A unified treatment*. Mahwah, NJ: Erlbaum.
- Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, 58(4), 525–543.
- Milfont, T. L., & Fischer, R. (2010). Testing measurement invariance across groups: Applications in cross-cultural research. *International Journal of Psychological Research*, 3(1), 111-121.
- OECD [Organisation for Economic Co-operation and Development] (2001). Knowledge and skills for life. First results from PISA 2000. Paris: OECD Publishing. Retrieved January 26, 2019, from <http://www.oecd.org/education/school/programmeforinternationalstudentassessmentpisa/33691620.pdf>.
- OECD [Organisation for Economic Co-operation and Development] (2014). TALIS 2013 results: An international perspective on teaching and learning. Paris: OECD Publishing. Retrieved July 25, 2018, from <http://dx.doi.org/10.1787/9789264196261-en>
- OECD [Organisation for Economic Co-operation and Development] (2016). PISA 2015 results (Volume I): Excellence and equity in education, PISA. Paris: OECD Publishing. Retrieved January 26, 2019, from <http://dx.doi.org/10.1787/9789264266490-en>.
- OECD [Organisation for Economic Co-operation and Development] (2017). *Education at a glance 2017. OECD indicators*. Paris: OECD Publishing. Retrieved July 25, 2018, from [http://www.oecd-ilibrary.org/education/education-at-a-glance-2017\\_eag-2017-en](http://www.oecd-ilibrary.org/education/education-at-a-glance-2017_eag-2017-en)
- Oishi, S. (2007). The application of structural equation modeling and item response theory to cross-cultural positive psychology research. In A. D. Ong & M. H. M. van Dulmen (Eds.), *Oxford handbook of methods in positive psychology* (pp. 126–138). New York: Oxford University Press.
- Papanastasiou, C., & Papanastasiou, E. (1997). Factors that influence students to become teachers. *Educational Research and Evaluation*, 3(4), 305–316.
- Paronen, P., & Lappi, O. (2018). Finnish teachers and principals in figures. Reports and Surveys 2018:4. Helsinki: Finnish National Agency for Education. Retrieved July 25, 2018, from [https://www.oph.fi/download/189802\\_finnish\\_teachers\\_and\\_principals\\_in\\_figures.pdf](https://www.oph.fi/download/189802_finnish_teachers_and_principals_in_figures.pdf)
- Perho, H., & Korhonen, M. (2012). Ammatillinen suuntautuminen, persoonallisuuspiirteet sekä työn laatu työuupumuksen, työn imun ja valintatyytyväisyyden tekijöinä. Lastentarhanopettajien 30 vuoden seuranta [Vocational orientation, personality traits and quality of work as factors of burnout, work engagement and career choice satisfaction: thirty-year follow-up of kindergarten teachers]. Reports and Studies in Education, Humanities, and Theology, No 6. Joensuu: University of Eastern Finland.
- Pyhältö, K., Pietarinen, J., & Salmela-Aro, K. (2011). Teacher–working-environment fit as a framework for burnout experienced by Finnish teachers. *Teaching and Teacher Education*, 27(7), 1101–1110.
- R Core Team (2017). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.

- Revelle, W. (2017). *psych: Procedures for psychological, psychometric, and personality research*. Evanston, Illinois: Northwestern University.
- Richardson, P. W., & Watt, H. M. (2006). Who chooses teaching and why? Profiling characteristics and motivations across three Australian universities. *Asia-Pacific Journal of Teacher Education*, 34(1), 27–56.
- Roness, D., & Smith, K. (2010). Stability in motivation during teacher education. *Journal of Education for Teaching*, 36(2), 169–185.
- Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48(2), 1–36.
- Rothland, M. (2013). Beruf: Lehrer/Lehrerin – Arbeitsplatz: Schule. Charakteristika der Arbeitstätigkeit und Bedingungen der Berufssituation [Profession: Teacher – Workplace: School. Work characteristics and job conditions]. In M. Rothland (Ed.), *Belastung und Beanspruchung im Lehrerberuf* (pp. 21–39). Wiesbaden: Springer VS.
- Rothland, M. (2014). Wer entscheidet sich für den Lehrerberuf? [Who chooses to become a teacher?] In E. Terhart, H. Bennewitz, & M. Rothland (Eds.), *Handbuch der Forschung zum Lehrerberuf* (pp. 319–348). Münster, New York: Waxmann.
- Sahlberg, P. (2011). *Finnish lessons: What can the world learn from educational change in Finland?* New York: Teachers College Press.
- Salifu, I., Alagbela, A. A., & Ofori, C. G. (2017). Factors influencing teaching as a career choice (FIT-Choice) in Ghana. *Teaching Education*, 29(2), 111–134.
- Sarja, A., Nyman, T., Ito, H., & Jaatinen, R. (2017). The foreign language teaching profession in Finnish and Japanese society – A sociocultural comparison. *Pedagogy, Culture & Society*, 25(2), 225–241.
- Sass, D. A. (2011). Testing measurement invariance and comparing latent factor means within a confirmatory factor analysis framework. *Journal of Psychoeducational Assessment*, 29(4), 347–363.
- Satorra, A., & Bentler, P. M. (1994). Corrections to test statistics and standard errors in covariance structure analysis. In A. von Eye & C. C. Clogg (Eds.), *Latent variables analysis: Applications to development research* (pp. 399–419). Thousand Oaks: SAGE Publications.
- Satorra, A., & Bentler, P. M. (2010). Ensuring positiveness of the scaled difference chi-square test statistic. *Psychometrika*, 75(2), 243–248.
- Schaarschmidt, U. (Ed.). (2005). *Halbtagsjobber? Psychische Gesundheit im Lehrerberuf – Analyse eines veränderungsbedürftigen Zustands* [Part-time jobber? Mental health in the teaching profession – Analysis of a situation in need of change]. Weinheim: Beltz.
- Statistics Finland (2016). *Women and men in Finland 2016*. Helsinki: Statistics Finland. Retrieved January, 28, 2019, from <https://www.stat.fi/uutinen/women-and-men-in-finland-2016>.
- Statistics Finland (2018). *Palkkatilastot* [Payment statistics]. Helsinki: Statistics Finland. Retrieved January 27, 2019, from <https://www.kt.fi/tilastot-ja-julkaisut/palkkatilastot>.

- Steenkamp, J.-B. E. M., & Baumgartner, H. (1998). Assessing measurement invariance in cross-national consumer research. *Journal of Consumer Research*, 25, 78–90.
- Steinmetz, H. (2011). Estimation and comparison of latent means across cultures. In E. Davidov, P. Schmidt, & J. Billiet (Eds.), *Cross-cultural analysis: Methods and applications* (pp. 85–116). New York: Routledge.
- Stöber, J. (2001). The social desirability scale-17 (SDS-17). *European Journal of Psychological Assessment*, 17(3), 222–232.
- Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston: Pearson.
- Van de Vijver, F. J. R., & Tanzer, N. K. (1997). Bias and equivalence in cross-cultural assessment: An overview. *European Review of Applied Psychology*, 47(4), 263–280.
- Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, 3(1), 4–70.
- Watt, H. M. G., & Richardson, P. W. (2007). Motivational factors influencing teaching as a career choice: Development and validation of the FIT-Choice scale. *Journal of Experimental Education*, 75, 167–202.
- Watt, H. M. G., & Richardson, P. W. (2012). An introduction to teaching motivations in different countries: Comparisons using the FIT-Choice scale. *Asia-Pacific Journal of Teacher Education*, 40(3), 185–197.
- Watt, H. M. G., Richardson, P. W., Klusmann, U., Kunter, M., Beyer, B., Trautwein, U., & Baumert, J. (2012). Motivations for choosing teaching as a career: An international comparison using the FIT-Choice scale. *Teaching and Teacher Education*, 28, 791–808.
- Watt, H. M. G., Richardson, P. W., & Smith, K. (Eds.) (2017). *Global Perspectives on Teacher Motivation*. Cambridge: Cambridge University Press.
- Weishaupt, H. (2014). Lehrerberuf: Arbeitsplatz, Arbeitszeit, Arbeitsmarkt [Teacher profession: Workplace, working time, employment market]. In E. Terhart, H. Bennewitz, & M. Rothland (Eds.), *Handbuch der Forschung zum Lehrerberuf* (pp. 103–122). Münster, New York: Waxmann.
- Wigfield, A., & Eccles, J. S. (2000). Expectancy–value theory of achievement motivation. *Contemporary Educational Psychology*, 25(1), 68–81.

*Appendix A. FIT-Choice motivation subscales including English and Finnish items.*

<b>Factor</b>	<b>Item*</b>	<b>Original English</b>	<b>Finnish translation</b>
<i>Perceived teaching abilities</i>	B5	I have the qualities of a good teacher.	Minulla on hyvän opettajan ominaisuudet.
	B19	I have good teaching skills.	Minulla on hyvät opetustaidot.
	B43	Teaching is a career suited to my abilities.	Opettajan ura sopii kykyihini.
<i>Intrinsic value</i>	B1	I am interested in teaching.	Olen kiinnostunut opettamisesta.
	B7 <sup>†</sup>	I have always wanted to be a teacher.	Olen aina halunnut opettajaksi.
	B12	I like teaching.	Pidän opettamisesta.
<b>Personal utility values</b>			
<i>Job security</i>	B14	Teaching will offer a steady career path.	Opetustyö tarjoaa vakaan urapolun.
	B27	Teaching will provide a reliable income.	Opetustyö tarjoaa luotettavat ansiotulot.
	B38	Teaching will be a secure job.	Opetusalalla työpaikka on varma.
<i>Time for family</i>	B2	Part-time teaching could allow more family time.	Osa-aikainen opetustyö voisi suoda enemmän aikaa perheelle.
	B16	Teaching hours will fit with the responsibilities of having a family.	Opetustunnit sopivat yhteen perheellisen velvoitteiden kanssa.
	B29	School holidays will fit in with family commitments.	Koulun lommat sopivat yhteen perhesitoumusten kanssa.
	B4 <sup>†</sup>	As a teacher I will have lengthy holidays.	Opettajana minulla tulee olemaan pitkäkköt lommat.
	B18	As a teacher I will have a short working day.	Opettajana minulla tulee olemaan lyhyet työpäivät.
<b>Social utility values</b>			
<i>Shape future of children/adolescents</i>	B9	Teaching will allow me to shape child/adolescent values.	Opetustyössä voin muokata lasten/nuorten arvoja.
	B23	Teaching will allow me to influence the next generation.	Opetustyössä voin vaikuttaa seuraavaan sukupolveen.
<i>Enhance social equity</i>	B36	Teaching will allow me to raise the ambitions of under-privileged youth.	Opetustyössä voin saada heikommassa asemassa olevat nuoret asettamaan tavoitteensa korkeammalle.
	B49	Teaching will allow me to benefit the socially disadvantaged.	Opetustyössä voin olla hyödyksi sosiaalisesti heikommassa asemassa oleville.
<i>Make social contribution</i>	B6	Teaching will allow me to provide a service to society.	Opetustyössä voin palvella yhteiskuntaa.
	B20	Teachers make a worthwhile social contribution.	Opettajien yhteiskunnallinen panos on merkittävä.
	B31 <sup>†</sup>	Teaching enables me to give back to society.	Opetustyössä voin maksaa takaisin yhteiskunnalle.

<b>Factor</b>	<b>Item*</b>	<b>Original English</b>	<b>Finnish translation</b>
<i>Work with children/adolescents</i>	B10 <sup>†</sup>	I want to help children and adolescents learn.	Haluan auttaa lapsia ja nuoria oppimaan.
	B13	I want a job that involves working with children and adolescents.	Haluan työn, jossa toimitaan lasten/nuorten parissa.
	B26	I want to work in a child/adolescent-centred environment.	Haluan työskennellä ympäristössä, jossa lapset/nuoret ovat keskiössä.
	B37	I like working with children/adolescents.	Pidän työskentelystä lasten/nuorten parissa.
<i>Fallback career</i>	B11	I was unsure of what career I wanted.	Olin epävarma siitä, mille uralle haluaisin.
	B35	I was not accepted into my first-choice career.	Minua ei hyväksytty ykkösvaihtoehtonani olleelle alalle.
	B48	I chose teaching as a last-resort career.	Valitsin opetustyön viimeisenä uravaihtoehtona.
<i>Prior teaching and learning experiences</i>	B17	I have had inspirational teachers.	Minulla on ollut inspiroivia opettajia.
	B30	I have had good teachers as role models.	Minulla on ollut hyviä opettajia roolimalleina.
	B39	I have had positive learning experiences.	Minulla on ollut myönteisiä oppimiskokemuksia.
<i>Social influences</i>	B3	My friends think I should become a teacher.	Ystäväni mielestä minusta pitäisi tulla opettaja.
	B24	My family thinks I should become a teacher.	Perheeni mielestä minusta pitäisi tulla opettaja.
	B40	People I have worked with think I should become a teacher.	Niiden mielestä, joiden kanssa olen työskennellyt, minusta pitäisi tulla opettaja.

*Note.* All motivation items were listed in a single block introduced by the phrase: “I chose to become a teacher because...” (Finnish; “Päätin opiskella opettajaksi, koska...”). Each item could be answered on a 7-point Likert scale ranging from 1 = “not at all important” to 7 = “extremely important” (Finnish: 1 = “Ei lainkaan tärkeä” to 7 = “Erittäin tärkeä”). \* The original item code used by Watt and Richardson (2007) was used here. <sup>†</sup> Items that had to be removed from the analysis based on insufficient psychometric properties.

*Appendix B. FIT-Choice perception subscales including English and Finnish items.*

<b>Factor</b>	<b>Item*</b>	<b>Original English</b>	<b>Finnish translation</b>
<b>Task demand</b>			
<i>High demand</i>	C2	Do you think teachers have a heavy workload?	Onko opettajilla mielestäsi raskas työtaakka?
	C7	Do you think teaching is emotionally demanding?	Onko opetustyö mielestäsi emotionaalisesti vaativaa?
	C11	Do you think teaching is hard work?	Onko opettaminen mielestäsi kovaa työtä?
<i>Expert career</i>	C10	Do you think teaching requires high levels of expert knowledge?	Vaatiiko opetustyö mielestäsi paljon asiantuntijatietaoa?
	C14	Do you think teachers need high levels of technical knowledge?	Tarvitsevatko opettajat mielestäsi paljon teknistä tietoa?
<b>Task return</b>			
<i>Social status</i>	C4	Do you believe teachers are perceived as professionals?	Nähdäänkö opettajat mielestäsi ammattilaisina?
	C8	Do you believe teaching is perceived as a high-status occupation?	Nähdäänkö opetustyö mielestäsi korkean statuksen ammattina?
	C12	Do you believe teaching is a well-respected career?	Onko opettaminen mielestäsi urana arvostettu?
	C9	Do you think teachers feel valued by society?	Tuntevatko opettajat mielestäsi olevansa yhteiskunnan arvostamia?
	C13	Do you think teachers feel their occupation has high social status?	Tuntevatko opettajat mielestäsi ammattinsa sosiaalisen statuksen korkeaksi?
<i>Salary</i>	C1	Do you think teaching is well paid?	Onko opetustyö mielestäsi hyvin palkattua?
	C3	Do you think teachers earn a good salary?	Saavatko opettajat mielestäsi hyvää palkkaa?
<i>Social dissuasion</i>	D2	Were you encouraged to pursue careers other than teaching?	Kannustettiin sinua ryhtymään muulle kuin opetuslalle?
	D4	Did others tell you teaching was not a good career choice?	Sanottiin sinulle, ettei opettaminen ollut hyvä uravalinta?
	D6	Did others influence you to consider careers other than teaching?	Saivatko toiset sinut harkitsemaan jotain muuta kuin opettajan uraa?
<i>Satisfaction with choice</i>	D3 <sup>†</sup>	How satisfied are you with your choice of becoming a teacher?	Kuinka tyytyväinen olet valintaasi hakeutua opettajaksi?
	D5 <sup>†</sup>	How happy are you with your decision to become a teacher?	Kuinka onnellinen olet päätöksestäsi ryhtyä opettajaksi?

*Note.* The items of the task demand and task return subscales were introduced by the phrase, “For each question below, please rate the extent to which you agree it is true about teaching” and followed by the question, “Do you think that...?” (Finnish: “Arvioi, missä määrin seuraavat kysymykset pitävät omalta osaltasi paikkansa.”). The social dissuasion and satisfaction with choice items were introduced by the phrase, “For each question below,

please rate the extent to which it is true for you” (Finnish: “Arvioi, missä määrin seuraavat kysymykset mielestäsi pitävät paikkansa.”). Each item could be answered on a 7-point Likert scale ranging from 1 = “not at all” to 7 = “extremely” (Finnish: 1 = “Ei lainkaan” to 7 = “Mitä suurimmassa määrin”). \* The original item code used by Watt and Richardson (2007) was used here. † Items that had to be removed from the analysis based on insufficient psychometric properties.

### *Appendix C. Latent correlations between motivation factors.*

<b>Factors</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>	<b>9</b>	<b>10</b>	<b>11</b>
<i>1. Perceived teaching abilities</i>	–	<b>.68</b>	<b>.33</b>	.15	<b>.44</b>	<b>.27</b>	<b>.32</b>	<b>.20</b>	-.33	.02	<b>.29</b>
<i>2. Intrinsic value</i>	<b>.88</b>	–	.24	-.10	.19	<b>.34</b>	<b>.34</b>	<b>.36</b>	<b>-.46</b>	<b>.23</b>	<b>.20</b>
<i>3. Job security</i>	-.13	-.04	–	<b>.42</b>	.10	.19	<b>.29</b>	.01	.03	.11	.15
<i>4. Time for family</i>	-.14	<b>-.18</b>	<b>.58</b>	–	.09	-.08	-.14	-.13	<b>.27</b>	.00	<b>.24</b>
<i>5. Shape future of children/adolescents</i>	<b>.34</b>	<b>.30</b>	.06	<b>.21</b>	–	<b>.64</b>	<b>.59</b>	<b>.30</b>	.01	.05	.15
<i>6. Enhance social equity</i>	<b>.29</b>	<b>.29</b>	.07	<b>.23</b>	<b>.84</b>	–	<b>.56</b>	<b>.43</b>	-.09	.17	.12
<i>7. Make social contribution</i>	<b>.25</b>	<b>.36</b>	.07	.14	<b>.76</b>	<b>.89</b>	–	.22	-.10	<b>.23</b>	.14
<i>8. Work with children/adolescents</i>	<b>.39</b>	<b>.51</b>	-.08	.01	<b>.24</b>	<b>.37</b>	<b>.30</b>	–	-.29	.02	<b>.19</b>
<i>9. Fallback career</i>	<b>-.46</b>	<b>-.57</b>	<b>.25</b>	<b>.27</b>	-.06	.02	-.08	<b>-.37</b>	–	.04	-.08
<i>10. Prior teaching and learning experiences</i>	.12	<b>.20</b>	.07	.09	<b>.31</b>	.17	<b>.23</b>	.02	-.04	–	<b>.21</b>
<i>11. Social influences</i>	<b>.24</b>	<b>.20</b>	.06	.14	<b>.26</b>	<b>.27</b>	<b>.36</b>	<b>.20</b>	-.03	<b>.30</b>	–

*Note.* Product-moment correlations for the Finnish sample above the diagonal. Product-moment correlations for the German sample below the diagonal. Correlations in bold are significant at the 5% level (two-tailed test).

### *Appendix D. Latent correlations between perception factors.*

<b>Factors</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
<i>1. High demand</i>	–	<b>.55</b>	-.07	<b>-.26</b>	.07
<i>2. Expert career</i>	<b>.56</b>	–	.09	-.16	-.02
<i>3. Social status</i>	.14	<b>.33</b>	–	<b>.46</b>	<b>-.25</b>
<i>4. Salary</i>	<b>.22</b>	<b>.26</b>	<b>.58</b>	–	-.11
<i>5. Social dissuasion</i>	.00	-.06	-.19	-.07	–

*Note.* Product-moment correlations for the Finnish sample above the diagonal. Product-moment correlations for the German sample below the diagonal. Correlations in bold are significant at the 5% level (two-tailed test).