Author(s): Yada, Akie; Tolvanen, Asko; Malinen, Olli-Pekka; Imai-Matsumura, Kyoko; Shimada, Hiroshi; Koike, Rihei; Savolainen, Hannu

Title: Teachers' self-efficacy and the sources of efficacy : A cross-cultural investigation in Japan and Finland

Year: 2019

Version: Accepted version (Final draft)

Please cite the original version:

Yada, A., Tolvanen, A., Malinen, O.-P., Imai-Matsumura, K., Shimada, H., Koike, R., & Savolainen, H. (2019). Teachers' self-efficacy and the sources of efficacy : A cross-cultural investigation in Japan and Finland. Teaching and Teacher Education, 81, 13-24. https://doi.org/10.1016/j.tate.2019.01.014

Teachers' self-efficacy and the sources of efficacy: A cross-cultural investigation in Japan and Finland

Akie Yada[a], Asko Tolvanen[b], Olli-Pekka Malinen[c,d], Kyoko Imai-Matsumura[e], Hiroshi Shimada[e], Rihei Koike[f], Hannu Savolainen[a,g]

[a] *Faculty of Education and Psychology, Department of Education, University of Jyvaskyla, P.O. Box 35 (Ruusupuisto), 40014, Jyväskylä, Finland*

[b] *Faculty of Education and Psychology, Methodology Centre for Human Sciences, University of Jyväskylä, P.O. Box 35 (Ruusupuisto), 40014, Jyväskylä, Finland*

[c] *Niilo Mäki Institute, P.O. Box 35, 40014, Jyväskylä, Finland*

[d] *EduCluster Finland, affiliate of the University of Jyväskylä, Lutakonaukio 7, 40100 Jyväskylä, Finland*

[e] *Graduate School of Education, Hyogo University of Teacher Education, 942-1 Shimokume, Kato, Hyogo, 673-1494, Japan*

[f] *Himeji Municipal Shikama-chubu Junior High School, 206 Shikamaku Hosoe, Himeji, Hyogo, 672-8064, Japan*

[g] *Optentia Research Focus Area, North West University, Vaal Triangle Campus, P.O. Box 1174, Vanderbijlpark 1900, South Africa*

Corresponding author: ado.eika@gmail.com (A. Yada)

ABSTRACT

The study explores the extent and sources of Teachers' Self-Efficacy (TSE) for inclusive practices among 261 Japanese and 1123 Finnish teachers. Measurement invariance was tested to ensure the chosen scales' cross-cultural validity. In both countries, mastery experience was identified as the strongest of the four sources contributing uniquely to TSE. However, the two groups differed in how verbal persuasion predicted TSE. The findings indicate that the effects of the four sources on TSE depend strongly on sociocultural context, and that, in Japan, other sources may exert a powerful influence. Practical implications are discussed, with particular regard to teacher training programs.

Keywords: Teachers' self-efficacy; sources of self-efficacy; inclusive education; cross-cultural measurement invariance, multi-trait multi-method model, Cholesky decomposition approach

Highlights:

- Teachers' self-efficacy (TSE) and its sources were examined in Japan and Finland.
- Measurement invariance testing confirmed the scales' construct validity.
- Mastery experience made the strongest unique contribution to TSE in both countries.
- The influence of verbal persuasion on TSE differed between Japan and Finland.
- Sources other than those proposed by Bandura (1977) may influence TSE in Japan.

1. Introduction

Since the publication of the Salamanca Statement and Framework for Action on Special Needs Education (UNESCO, 1994), inclusive education has gradually entered the mainstream around the world, reinforcing a global agenda to offer equal educational opportunities to all children (United Nations, 2006; United Nations General Assembly, 2015). Teachers clearly play an important role in implementing inclusive education, and a number of studies have highlighted factors associated with teachers who create an inclusive classroom environment (e.g., Avramidis & Norwich, 2002; de Boer, Jan Pijl, & Minnaert, 2011). One such factor that has attracted research interest is Teachers' Self-Efficacy (TSE) for inclusive practices (e.g., Malinen, Savolainen, & Xu, 2013; Meijer & Foster, 1988; Soodak & Podell, 1993). For instance, Meijer and Foster (1988) suggested that teachers' higher self-efficacy scores were associated with lower ratings of student problem seriousness—that is, whether a pupil is seen to be causing a significant problem in their classroom. TSE was also related to whether teachers thought it better to refer such pupils to special education (Meijer & Foster, 1988). Some studies of the sources of TSE refer to Bandura's (1997) theory (see Morris, Usher, & Chen, 2017), but while there has been extensive research on TSE for inclusive practices, much less is known about the sources of TSE in this context.

A substantial body of recent literature emphasizes the need for cross-cultural studies of TSE, which is often contingent on cultural and historical background (e.g., Chiu & Klassen, 2009; Klassen, 2004b; Markus & Kitayama, 1991; Mitchell, 2005). Markus and Kitayama (1991) demonstrated that current theories about the understanding of self and others refer principally to Western cultural contexts and noted that the situation may differ in non-Western contexts. As the development of self-efficacy entails understanding oneself in relation to others, how the sources of self-efficacy affect TSE for inclusive practices is also likely to depend on cultural background. As very little is known about the sources of TSE in different cultural historical settings or in different countries, the purpose of this paper is to explore the relationship between TSE for inclusive practices and its sources in one non-Western country in East Asia (Japan) and one Western country in Nordic region (Finland).

1.1. TSE and its sources

Bandura (1977) introduced the term *self-efficacy* as an element of his social cognitive theory, defining it as the belief that one can perform effectively in a given situation. TSE is specific to teachers, which can be understood as a teacher's beliefs about their ability to promote student learning (Klassen, Tze, Betts, & Gordon, 2011; Ross & Bruce, 2007). This is assumed to relate both to teachers' behavior and affect (e.g., adopting new teaching strategies, burnout, and stress) and to student outcomes such as academic achievement, motivation, and efficacy (Klassen et al., 2011; Ross & Bruce, 2007; Tschannen-Moran, Hoy, & Hoy, 1998).

Various studies have assessed TSE for teaching with different academic domains such as teaching math, physical education or language (Klassen et al., 2011). One of the domains on which researchers' attention has focused in last decade is TSE for inclusive practices. It has been suggested that teachers with higher sense of TSE for inclusive practices are more willing to teach students with special needs in their classrooms (e.g., Meijer & Foster, 1988; Savolainen, Engelbrecht, Nel, & Malinen, 2012). Futhermore, several variables (e.g., experience in teaching student with special needs, amount of inclusive education training and knowledge of legislation and policy) have been found to predict TSE for inclusive practices (Yada, Tolvanen, & Savolainen, 2018; Forlin, Sharma, & Loreman, 2014). A number of authors have also considered the effects of socio-cultural contexts on TSE for inclusive practices and the similarities and differences were discussed using cultural-historical and legal frameworks (e.g., Savolainen et al., 2012; Yada et al., 2018; Sharma, Aiello, Pace, Round, & Subban, 2018). For instance, the result of a study (Yada & Savolainen, 2017) indicated that TSE for inclusive practices was relatively low in Japan. The result was explained from the perspective of Japanese educational system (e.g., inadequate training for teachers in inclusive practices) and its culture (e.g., Japanese people's disposition to be modest).

Bandura (1997) proposed that there are four sources of self-efficacy. The first of these is mastery experience—that is, experience of success or failure in a specific situation. How this affects self-efficacy will depend on the process and on the effort made to overcome obstacles. Self-efficacy is higher when individuals frame their past accomplishments in a positive way (Chen & Usher, 2013). Previous studies have supported Bandura's view that mastery experience is the most

powerful of the four sources of self-efficacy (Tschannen-Moran & Hoy, 2007; Usher & Pajares, 2008). The second source is vicarious experience based on modeling the attainments of others, and group norms and one's relationship with others can enhance or diminish efficacy beliefs. Vicarious experience has been shown to exert a more powerful influence when the model is perceived as similar in terms of ability and/or personal characteristics such as age, gender, and ethnicity (Bandura, 1997; Usher & Pajares, 2008). Vicarious experience plays a fundamental role in situations where one is given a new task for which the criteria of proficiency are unclear (Bandura, 1997; Chen & Usher, 2013). The third source of self-efficacy is verbal persuasion, which can be defined as appraisal or evaluative feedback from others. Although verbal persuasion alone is less powerful than the two preceding sources (Bandura, 1997), it can improve efficacy beliefs where positive and sincere evaluation realistically reflects the agent's capabilities (Schunk, 1984). The opposite is also true; devaluative feedback can undermine self-efficacy (Hattie & Timperley, 2007). Bandura's fourth source of self-efficacy is psychological and affective state. When people judge their capabilities, they sometimes utilize somatic information—for example, higher stress levels or negative emotional proclivities can undermine perceived self-efficacy (Bandura, 1997). In their review of the literature, Morris et al. (2017) pointed out that negative psychological and affective states are more often investigated in this context, even though positive states (e.g., feelings of excitement) can contribute to enhanced self-efficacy (Mills, 2011).

Several variables such as gender or ethnicity are known to contribute to the influence of different sources of self-efficacy. Usher and Pajares (2008) demonstrated that gender differences in the influence of specific sources were often domain-specific; for example, while male students reported more mastery experience in the area of science (Britner & Pajares, 2006), female students showed greater mastery experience in writing (Pajares, Johnson, & Usher, 2007). To study the role of ethnicity, Stevens, Olivárez, and Hamman (2006) compared math self-efficacy and its sources among Hispanic and White students in the 4th to 10th grades. They found that Hispanic students more frequently mentioned access to good models (vicarious experience) as a source of efficacy, with fewer experiences of praise (verbal persuasion) and success (mastery experience) than among White students. In another study, Klassen (2004a) investigated the differential impact of the four sources on mathematics efficacy beliefs among grade 7 Indo-Canadian and Anglo-Canadian students. Vicarious experience and verbal persuasion were significant predictors of mathematics

efficacy in Indo-Canadian students but not among Anglo-Canadian students. The results may indicate that self-oriented sources (mastery experience and psychological and affective states) predominated in an individualist cultural group while other-oriented sources (vicarious experience and verbal persuasion) were stronger in a collectivist culture (Klassen, 2004a). Although the above results refer to students' self-efficacy, these findings may also be applicable to teachers.

Over the past decade, much more information has emerged in relation to sources of TSE. In a recent systematic literature review, Morris et al. (2017) found that over half of the studies meeting their search criteria were published between 2010 and 2015. Based on this literature, it is clear that mastery experience is the strongest source of TSE while the other three sources also exert a positive or negative influence (e.g., Bruce & Ross, 2008; Milner, 2002). Although many existing studies provide important insights into the sources of TSE, Morris et al. (2017) identified several problems and inconsistencies. First, they argue that few scales measure all four sources, and these scales are not psychometrically strong enough. For instance, Poulou's (2007) factor analysis of the sources of TSE among 198 student teachers revealed that the mastery experience and verbal persuasion factors did not separate psychometrically as expected. Morris et al. (2017) suggest that this is also the case in other studies because the four sources mediate and moderate each other's effect on TSE. Additionally, they contended that existing accounts have not dealt with the independent effect on TSE of each hypothesized source (Morris et al., 2017). In a professional development training program designed to influence the four sources of TSE information, Ross and Bruce (2007) found that the program had a positive effect on teachers' beliefs about their ability to manage mathematics classrooms. However, they failed to specify which source of TSE information contributed to the change in participants' efficacy.

Beyond the four sources of TSE, other factors may also contribute. For example, some authors (Morris et al., 2017; Palmer, 2011; Wheatley, 2005; Wyatt, 2014) have suggested that gaining sufficient knowledge in specific areas (e.g., pedagogical, technological and subject-matter knowledge), which Morris et al. (2017) refer to as "mastery of knowledge," may add some variation to TSE. Previous studies have also found that respect and confidence from students and parents can strengthen TSE (Cheung, 2008; Milner, 2002; Milner & Hoy, 2003). Similarly, a sense of collective efficacy, which means teachers' shared perception that the school faculty as a whole is able to

produce positive effect on their students by organizing and executing the courses of action, can increase or decrease TSE (Goddard & Goddard, 2001). There may be further sources beyond those referred to here, and there is ongoing discussion as to whether those other factors are independent or form part of the four known sources (e.g., mastery of knowledge may form part of mastery experience) (Morris et al., 2017).

1.2. Cultural context for inclusive education in Japan and Finland

Both the Japanese and Finnish governments have promoted inclusive education in line with the Salamanca Statement (UNESCO, 1994), the Convention on the Rights of Persons with Disabilities (United Nations, 2006), and the 2030 Agenda for Sustainable Development (United Nations General Assembly, 2015). However, the two countries have adopted different approaches based on their unique historical and cultural background. Although the general concept of inclusive education encompasses groups such as children from ethnic minorities, low socioeconomic groups, or otherwise disadvantaged backgrounds (Mitchell, 2005), the Japanese inclusive education model focuses more on supporting children with disabilities (Forlin, Kawai, & Higuchi, 2015; MEXT, 2012). This may be because Japan's highly homogenous society includes fewer immigrants or refugees than other countries (OECD, 2018; Smith, Bond, & Kâğıtçıbaşı, 2006). In addition, it seems that the Japanese education system still relies on the medical model of disability (Ichikawa, 2016), and children who cannot accommodate the demands of mainstream schooling norms are likely to receive "custodial forms of care" (Borovoy, 2008) in separate special schools or classes. In light of the present situation, MEXT (2012) proposed that as many children as possible, regardless of disabilities, should study in regular schools, with special education schools serving as centers to support children, parents, teachers, and school staff and to build community networks. Yet, although the government has set this goal for inclusive education, deficits in appropriate teacher knowledge (Fujii, 2014; Ueno & Nakamura, 2011), pre- and in-service teacher training (Forlin et al., 2015), and collaborative work in schools (Ogiso & Tsuzuki, 2016) continue to challenge Japan's implementation of inclusive education.

In Finland, the latest significant reforms in relation to inclusive education began with the Strategy of Special Education (Ministry of Education of Finland, 2007). The multi-tiered system of

support called “Learning and Schooling Support” was adopted following the Act for Amendment of Basic Education Act in 2010 (Jahnukainen & Itkonen, 2016). This support is mandatory in all schools and comprises three levels: general support, intensified support, and special support (Björn, Aro, Koponen, Fuchs, & Fuchs, 2016; Jahnukainen & Itkonen, 2016). Part-time special education, remedial teaching, and/or guidance are offered for all children who need them in tier 1 (general support). Administrative decisions are required only in tier 3 (special support) if children need long-term support (FNBE, 2016). While the Finnish education system seems to succeed in offering flexible and equal education for every child, some critics have argued that some challenges remain. For example, it has been noted that some students regarded as incapable of attending regular classrooms are instead placed in self-contained special classes or schools, even though tier 3 support can also be organized in full-time inclusive settings through an individual education plan (Jahnukainen, 2011; Kivirauma, Klemelä, & Rinne, 2006). These “segregated tracks” continue beyond post-compulsory education into the individual’s adult life (Hakala, Björnsdóttir, Lappalainen, Jóhannesson, & Teittinen, 2018). Another concern is that although municipalities and schools must comply with the Act, authority to organize how special education is delegated to each municipality and school, giving them considerable autonomy in formulating and implementing school curricula (Pesonen et al., 2015). Consequently, special education philosophy and implementation strategies are seen to vary by municipality and even by school (Pesonen et al., 2015).

1.3. Validity of cross-cultural research

In a recent review of the TSE literature, Klassen et al. (2011) concluded that further investigation is essential in different cultural settings, especially in non-North American contexts, to strengthen the validity and generalizability of TSE theory. While cross-cultural research helps to identify interesting similarities and differences across countries, there are some challenges in ensuring valid comparison of different groups. A first major drawback of this approach is that educational concepts such as “inclusive education” or “self-efficacy” may be differently understood in different countries, even when using the same research instruments (Mitchell, 2005). Second, as one element of sociocultural divergence, linguistic differences may affect participants’ responses (Jahnukainen, 2015). Finally, cultural differences such as individualism or collectivism may influence both study

results and participants' response style. Several theories of cultural dimensions have been proposed (e.g., Hall, 1976; Hofstede, 2001; Schwartz, 1999); more specifically, previous studies have explored how living in an individualist or collectivist culture influences definitions of "self" (Klassen, 2004b; Markus & Kitayama, 1991). An individualist culture is characterized by an emphasis on "I" consciousness and the independence of groups to which a person belongs (e.g., family, organization, nation). In collectivist culture, on the other hand, high value is placed on "we" consciousness and group interdependence (Hofstede, 2001; Markus & Kitayama, 1991).

The additional perspective of horizontal and vertical cultures has further enhanced understanding of the different kinds of individualism and collectivism (Triandis, 2001). While a horizontal culture emphasizes equality, a vertical culture is characterized by hierarchy; together, these generate the four dimensions of horizontal individualism, horizontal collectivism, vertical individualism, and vertical collectivism (Triandis, 2001). Of the two countries in this study, Japan can be characterized as a vertical collectivist culture, in which people regard group superiority as important (Spielberger, 2004). Finland can be understood as a culture of horizontal individualism, emphasizing the equality of all people and each person's uniqueness (Triandis, 2001). As one example of the possible cultural contingency of participants' response style, "modesty bias" is assigned greater weight in the collectivist culture of East Asian countries (including Japan), where it is preferable to present oneself as average within a group. This is likely to result in lower scores on such measures as self-esteem and self-efficacy (Kagitçibasi, 1997; Markus & Kitayama, 1991; Vieluf, Kunter, & van de Vijver, 2013), even though Finnish people consider modesty to be one of their national virtues (Nishimura, Nevgi, & Tella, 2008).

Although it is impossible to completely eliminate such influences, measurement invariance is commonly tested to determine whether the same constructs are being measured in independent groups (Chen, Sousa, & West, 2005). While some TSE studies have tested cross-cultural measurement invariance (e.g., Brouwers & Tomic, 2001; Klassen et al., 2009; Yada et al., 2018), measurement invariance has not yet been investigated in relation to the sources of TSE in different countries.

1.4. Research questions

While a number of previous studies have investigated TSE for inclusive practices, only a few have examined the sources of TSE for inclusive practices, and still fewer have compared the sources of TSE across different cultural contexts. The aim of the present study was to measure all four sources of TSE for inclusive practices in two countries with differing cultural and historical backgrounds. In addition, this is the first study to use a psychometrically developed scale to examine the relationships between the four sources and TSE for inclusive practices, and in particular whether each source contributes uniquely to TSE. To that end, the study addressed the following research questions.

(1) Do the two scales used in this study measure the same constructs of TSE and sources of efficacy in both Japan and Finland?

(2) How do the four sources of self-efficacy predict TSE, and what is their individual contribution in predicting TSE in Japan and Finland?

Based on previous findings (Bandura, 1997; Tschannen-Moran & Hoy, 2007; Usher & Pajares, 2008; Klassen, 2004a), we formulated the following hypotheses in relation to the second research question.

Hypothesis 1: Mastery experience is the most influential source of TSE for inclusive practices in both Japan and Finland.

Hypothesis 2: The self-oriented sources of TSE (mastery experience and psychological and affective states) are more influential in Finland, which is a more individualist culture, while the other-oriented sources (vicarious experience and verbal persuasion) are more influential in Japan, which is a collectivist culture.

2. Method

2.1. Participants and procedure

The participants in the current study were in-service teachers working in primary and/or lower secondary schools in Japan and Finland. Teachers' participation was voluntary, and participants were informed by letter about the purposes of the research, data confidentiality, and their right to withdraw at any time. The schools represent a convenience sample of those that agreed to participate in the study.

The Japanese sample ($N = 261$) was collected from schools in western Japan in 2017. Hard copies of the questionnaire were distributed at each school, and the researcher visited the school to collect these on completion.

The Finnish sample ($N = 1123$) was drawn in the first phase of the ProKoulu project (2013–2014) from schools in the eastern part of Finland. The project, which ran from 2013 to 2016, investigated how school-wide positive behavior support works at school level. An online survey strategy was adopted for the Finnish component. Details of participants from both countries are presented in Table 1. In both cases, gender ratio and mean age were close to those of the general population of teachers (MEXT, 2017; OECD, 2013). Regarding to gender ratio, 11.0% in Finland and 0.4% in Japan had missing data in this variable.

Table 1. Participant background information

	Japan	Finland
Gender	Female 60.5%	Female 65.9%
	Male 39.1%	Male 23.1%
Mean age (*SD*)	39.82 (11.49)	45.19 (9.43)
School type (Grade)	Primary school (1–6) 57.5%	Primary (1–6) or comprehensive (1–9) school 65.0%

	Lower secondary school (7–9) 42.5%	Lower secondary school (7–9) 35.0%

2.2. Measures

TSE for inclusive practices was measured using the 18-item Teacher Efficacy for Inclusive Practices (TEIP) scale (Sharma, Loreman, & Forlin, 2012), which was developed specifically to measure this construct. Six items measure participants' *efficacy in instruction* (e.g., "I am confident in designing learning tasks so that the individual needs of students with disabilities are accommodated"). Six further items measure *efficacy in collaboration* (e.g., "I am confident in my ability to get parents involved in school activities of their children with disabilities" and "I am able to work jointly with other professionals and staff (e.g., aides and other teachers) to teach students with disabilities in the classroom"). The remaining six items measure *efficacy in managing behavior* (e.g., "I am able to calm a student who is disruptive or noisy"). TEIP items were originally scored from 1 (*Strongly Disagree*) to 6 (*Strongly Agree*), and the same Likert scale was adopted for the Japanese data. However, the Finnish data used a nine-point Likert scale ranging from 1 (*Not at all*) to 9 (*Very much*) and the reason for this transformation was to maintain consistency with other scales used in the larger study where the Finnish data of this study was taken from. Thus, to ensure that the both data were comparable, for the purposes of this study, we transposed TEIP scores from both data sets to match to a range between 0 and 1, where 0 indicates the lowest TEIP score and 1 indicates the highest. This was done by subtracting one from each score and dividing the result by five (for 6-point scores) or by eight (for 9-point scores)—that is, $(X\text{-}1)/(n\text{-}1)$. The TEIP scale was ealier validated in Japan by Yada and Savolainen (2017), and in Finland by Savolainen et al., (2012).

The sources of TSE were assessed using the Sources of Teacher Self-Efficacy (STSE) scale, which was developed as part of the ProKoulu project in Finland (Malinen, 2014). The scale comprises 16 items exploring the extent to which the four sources have affected participants' perceptions of their capabilities in each of the four teaching domains (instruction, behavior management, collaboration, and student engagement). The items can be divided into four subscales: a) mastery experience; b) vicarious experience; c) verbal persuasion; and d) affective state. A Likert-type scale ranging from 1 (*Not at all*) to 9 (*Very much*) was used in both countries.

The TEIP scale had already been translated into Japanese and Finnish in previous studies (Savolainen et al., 2012; Yada & Savolainen, 2017). As the STSE scale was originally written in Finnish, the researchers first translated it into English and then into Japanese. The Japanese version of the STSE scale was sent to translators for proofreading, and any changes were carefully discussed with the researchers.

2.3. Statistical analysis

Statistical analysis was performed using Mplus software (version 7) for Mac (Muthén & Muthén, 2012). Using the MLR estimator function, model parameters were estimated with robust standard error and scale corrected chi-square values using the maximum likelihood method of full information. Missing values accounted for 1.1% of the Japanese data and 1.3% of the Finnish data. The Missing At Random (MAR) option was applied to handle missing values where full information was utilized without imputing the missing values. As the likelihood ratio test is known to be sensitive to sample size (MacCallum, Browne, & Cai, 2006), model fit was evaluated using a two-index strategy (Hu & Bentler, 1999), in which a cutoff value close to .06 for Root Mean Square Error of Approximation (RMSEA) and .08 for Root Mean Squared Residual (SRMR) indicates a good fit of model. In addition, a Comparative Fit Index (CFI) close to .95 was used for reference.

Utilizing Structural Equation Modeling (SEM), the analysis consisted of two main stages. In the first stage, measurement invariance was tested for both scales to answer research question 1. In the first step, theoretically driven Multi-Group Confirmatory Factor Analysis (MGCFA) was performed for the TEIP and STSE scales. As the STSE scale's unique structure assesses the extent to which the four sources affect teachers' capabilities in the four teaching domains, items that belong to the same teaching domain show a high correlation when MGCFA is performed only for the source factors. To resolve this problem, we applied a Multi-Trait Multi-Method (MTMM) design (Campbell & Fiske, 1959), in which the four sources served as trait factors, and the four teaching domains were treated as method factors. As can be seen from the hypothesized model (Figure 1), each observed variable loaded onto both trait and method factors, and the correlations between trait and method factors were set to zero (Byrne, 2013). The model enabled partialing out of the covariance between the method factors, and only the variance related to source factors remained for

further analysis. Following implementation of the freely estimated models for both scales, some error covariances between items were added for the TEIP scale as suggested by modification indices to improve the model. In the second step, factor loadings were set as equal between groups to test metric invariance. Changes in RMSEA (ΔRMSEA) were used to evaluate invariance among different consecutive models. According to Chen (2007), a change of less than .015 in RMSEA indicates model invariance. In the third step, scalar invariance was tested by setting factor loadings and intercepts as equal across the two groups. Previous studies have shown that a second-order factor model is applicable to the TEIP scale because of the high correlations between primary factors (Malinen, Savolainen, & Xu, 2013; Yada, Tolvanen, & Savolainen, 2018). For that reason, the next step was a no-constraint second-order factor model, with *efficacy in instruction*, *efficacy in collaboration*, and *efficacy in managing behavior* as the lower order factors. In the final step, the factor loadings were set as equal for the second-order factors in order to determine whether there was metric invariance between groups.

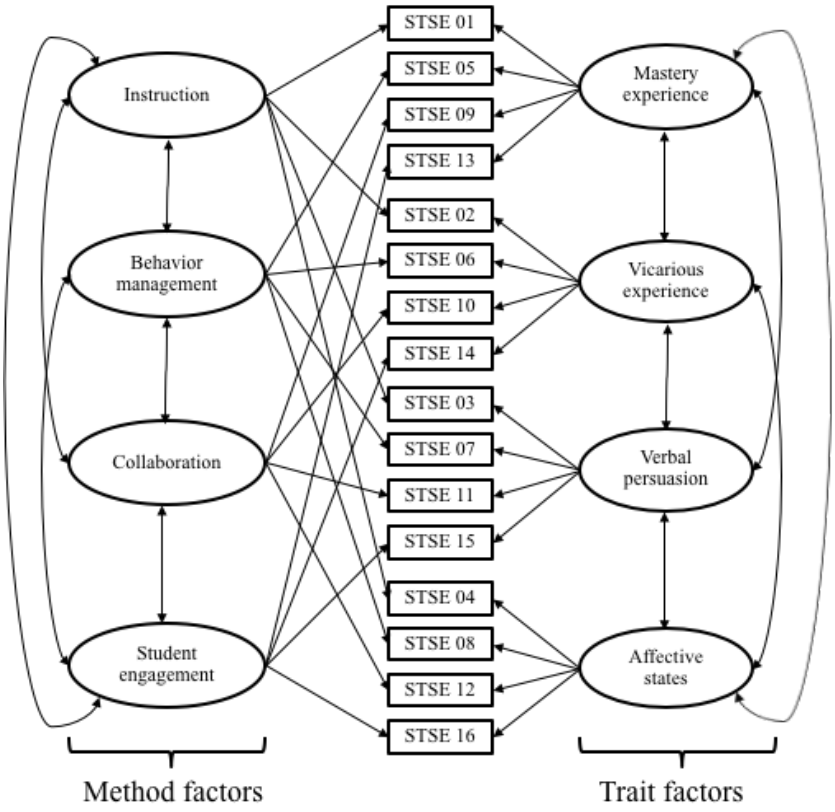

Figure 1. Hypothesized MTMM model for STSE scale

The second stage of analysis addressed research question 2; in this stage, the Japanese and Finnish data were analyzed separately. As the latent factors of the four sources were highly correlated, the Cholesky decomposition (de Jong, 1999) was employed to determine the unique contribution to TSE of those latent factors. This approach addresses the problem of multicollinearity utilizing a hierarchical regression analysis conducted in SEM (de Jong, 1999). More specifically, the four Cholesky factors partitioning the variance of the latent factors were entered into the regression model in a pre-determined order, and the Cholesky factor inserted lastly into the model represented the unique contribution of that factor to TSE. (See Figure 2 for an example of the Cholesky decomposition model where *affective states* was inserted lastly into the model.)

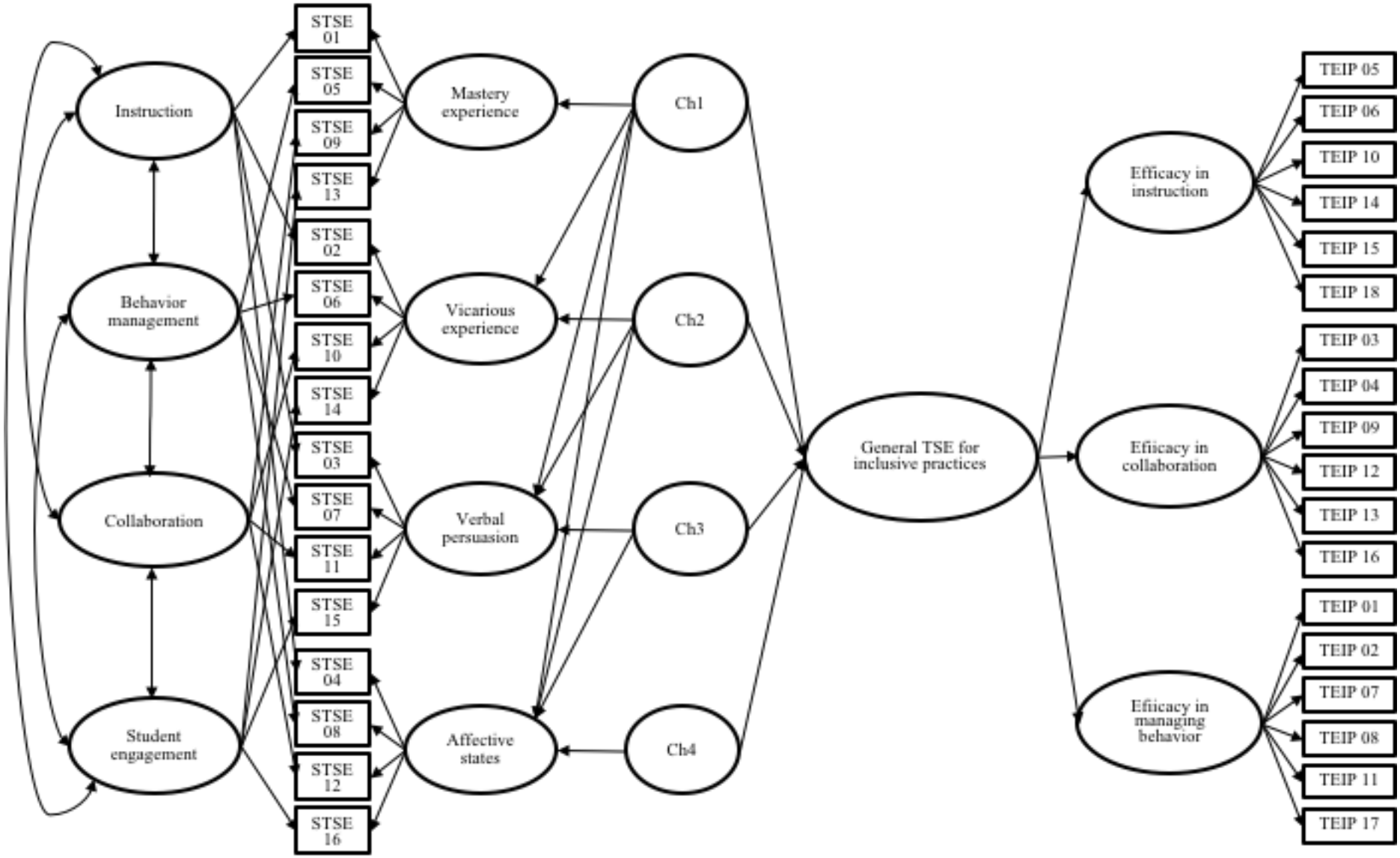

Figure 2. Example of the Cholesky decomposition model (order 1 in Tables 4 and 5)

3. Results

3.1. Testing measurement invariance

The first stage assessed the theoretically driven MGCFA model for the TEIP scale and the MTMM model for the STSE scale. All factor loadings for TEIP primary factors and STSE trait factors were statistically significant in both Japan and Finland. Table 2 below summarizes the results of model fit indices from less constrained to stricter model. First, the freely estimated model (Model 1) yielded a sufficient fit (RMSEA = .038; SRMR = .041; CFI = .957). Second, metric invariance was investigated (Model 2), and the model was found to exhibit adequate fit (RMSEA = .042; SRMR = .066; CFI = .946). The change in RMSEA between the no-constraint Model 1 and the constrained Model 2 was acceptable ($|\Delta\text{RMSEA}| = .004$). Third, to test scalar invariance, the factor loadings and intercepts were set as equal between groups. Model 3 showed acceptable fit (RMSEA = .046, SRMR = .072, and CFI = .933), with no great difference in RMSEA between this and the less constrained Model 2 ($|\Delta\text{RMSEA}| = .004$). The results indicate adequate invariance in the constructs, confirming the international validity of the two scales. As the three primary factors of the TEIP scale were highly correlated, the next step examined the second-order factor model. All three of the first-order factors had statistically significant factor loadings with the second-order factor in both groups. Based on previous studies (Malinen, Savolainen, & Xu, 2013; Yada, Tolvanen, & Savolainen, 2018), the second-order factor was named General Teacher Self-Efficacy for inclusive practices (GTSE). The freely estimated Model 4 with second-order factor achieved an acceptable fit (RMSEA = .047; SRMR = .073; CFI = .926), and there was no great change in RMSEA between Model 3 and Model 4 ($|\Delta\text{RMSEA}| = .001$). Next, factor loadings of the second-order factor were set as equal across countries. The metric invariance model of second-order factor (Model 5) achieved acceptable fit (RMSEA = .048; SRMR = .075; CFI = .925), supplemented by a change of .001 in RMSEA when compared with the less constrained Model 4. The results support metric invariance of second-order factors between Japan and Finland, indicating that the two scales used in this study measure the same constructs in both countries.

Table 2. Test of measurement invariance for the multi-group measurement model

Model	Explanation	Overall Fit Indices						Comparative Fit Indices	Model Comparison
		χ^2	df	*p*	RMSEA	SRMR	CFI	|ΔRMSEA|	
1	Freely estimated	1880.095	936	< .001	0.038	0.041	0.957	–	–
2	Factor loadings equal	2166.251	983	< .001	0.042	0.066	0.946	0.004	1 vs. 2
3	Factor loadings and intercepts equal	2461.590	1006	< .001	0.046	0.072	0.933	0.004	2 vs. 3
4	Factor loadings and intercepts equal for first-order factors Freely estimated for second-order factor	2657.631	1038	< .001	0.047	0.073	0.926	0.001	3 vs. 4
5	Factor loadings and intercepts equal for first-order factors Factor loadings equal for second-order factor	2667.706	1040	< .001	0.048	0.075	0.925	0.001	4 vs. 5

3.2. Comparing the effect of four sources of self-efficacy on TSE

In the second stage, Cholesky regression models were conducted separately in Japan and Finland to address the second research question. First, measurement models with all variables were analyzed. Standardized loadings of the first and second factors for the TEIP scale ranged from .56 to .98 for Japan and from .52 to .84 for Finland (all $p < .001$). The STSE scale trait factors showed statistically significant loadings ranging from .44 to .80 for Japan and from .63 to .84 for Finland (all $p < .001$). Table 3 shows correlations among the five latent factors in both countries. As correlations between the four factors of the STSE scales were medium to high, the Cholesky decomposition was used to avoid multicollinearity. Tables 4 and 5 below show the results of the Cholesky regression models for Japan (Table 4) and Finland (Table 5). The models provided an acceptable fit for both Japanese (RMSEA = .054; SRMR = .071; CFI = .929) and Finnish data (RMSEA = .038; SRMR = .043; CFI = .952).

The total R-squared values in Tables 4 and 5 indicate the extent to which the four sources of self-efficacy explain the variance in GTSE. The results show that the four sources of self-efficacy explained 54% of the variance in Finland but only 15% in Japan.

The first hierarchical regression models (Order 1 in Tables 4 and 5) investigated the unique contribution of "Affective States (AS)" on GTSE while controlling for the other three sources. "Mastery Experience (ME)" alone significantly predicted GTSE ($\beta = .36$, $p < .001$ for Japan and $\beta = .68$, $p < .001$ for Finland). Regarding the unique effect, AS did not account for additional variance in either Japan or Finland when the other three sources were controlled for.

The second hierarchical regression models (Order 2 in Tables 4 and 5) addressed the unique contribution of ME while controlling for the other three sources. The results indicate that "Vicarious Experience (VE)" alone predicted GTSE ($\beta = .23$, $p < .05$ for Japan and $\beta = .41$, $p < .001$ for Finland). In addition, ME accounted for unique variance in GTSE while the other three sources were controlled for ($\beta = .24$, $p < .01$ for Japan and $\beta = .35$, $p < .001$ for Finland), with a 6% increase in the explanation rate for Japan and a 12% increase for Finland.

The third hierarchical regression models (Order 3 in Tables 4 and 5) examined the unique contribution of VE while controlling for the other three sources. The results showed that "Verbal Persuasion (VP)" alone predicted GTSE in Finland ($\beta = .49$, $p < .001$) but not in Japan. There was no unique effect of VE on GTSE in either country.

The fourth hierarchical regression models (Order 4 in Tables 4 and 5) assessed the unique contribution of VP while controlling for the other three sources. AS alone predicted GTSE for both countries ($\beta = .24$, $p < .01$ for Japan and $\beta = .58$, $p < .001$ for Finland). Additionally, VP accounts for different degrees of unique variance in GTSE in Japan and Finland. For the Finnish sample, VP showed a unique positive contribution on GTSE ($\beta = .15$, $p < .001$) while the other three sources were controlled, yielding a 2% increase in explanation rate. On the other hand, the results for VP independent of the other three sources indicate a significant negative relationship between VP and GTSE ($\beta = -.14$, $p < .05$) in Japan, yielding a 2% increase in explanation rate. In other words, while VP alone was not associated with GTSE, higher VP scores predicted lower GTSE in Japan when ME, VE, and AS were taken into account. Lubin (1957) broadly explained Horst's (1941) definition of suppressor variable as subtracting some variance from a predictor, usually having a positive correlation with the predictor and zero correlation with a dependent variable. As these conditions were fulfilled, the results can be understood as a suppression effect but may also have happened by chance according to MacKinnon, Krull, and Lockwood (2000). For that reason, the results must be interpreted with caution.

In summary, the results show that ME had the strongest independent relationship with TSE in both countries, supporting Hypothesis 1. As there was no unique contribution from AS and VE in either country, Hypothesis 2 was not supported. However, the results suggest a possible difference between the two countries in terms of how VP affects GTSE.

Table 3. Correlations of latent factors

	1	2	3	4	5
1. Mastery experience (ME)	-	0.561***	0.530***	0.626***	0.360***
2. Vicarious experience (VE)	0.350***	-	0.837***	0.502***	0.227*
3. Verbal persuasion (VP)	0.369***	0.718***	-	0.553***	0.137
4. Affective states (AS)	0.720***	0.349***	0.460***	-	0.236**
5. GTSE	0.683***	0.409***	0.491***	0.593***	-

Notes: * $p < .05$, ** $p < .01$, *** $p < .001$. Correlations from the Japanese data are in the upper diagonal; correlations from the Finnish data are in the lower diagonal.

Table 4. Hierarchical regression analysis predicting GTSE in Japan (n = 261)

		Hierarchical Regression Analysis											
SIP	β	Order 1	β	ΔR^2	Order 2	β	ΔR^2	Order 3	β	ΔR^2	Order 4	β	ΔR^2
Mastery experience (ME)	0.34***	ME	0.36***	0.13	VE	0.23*	0.05	VP	0.14	0.02	AS	0.24**	0.06
Vicarious experience (VE)	0.24	VE	0.03	0.00	VP	-0.10	0.01	AS	0.19*	0.04	ME	0.27***	0.07
Verbal persuasion (VP)	-0.27*	VP	-0.14	0.02	AS	0.18*	0.03	ME	0.28***	0.08	VE	0.03	0.00
Affective states (AS)	0.06	**AS**	**0.04**	**0.00**	**ME**	**0.24****	**0.06**	**VE**	**0.12**	**0.02**	**VP**	**-0.14***	**0.02**
Total R^2	0.15			0.15			0.15			0.15			0.15

Notes: * $p < .05$, ** $p < .01$, *** $p < .001$. SIP = simultaneous inclusion of predictors in the regression model; $\Delta R2$ = incremental proportion of variance described in GTSE. The variables inserted lastly into the models and making a unique contribution are shown in bold.

Table 5. Hierarchical regression analysis predicting GTSE in Finland (n = 1123)

		Hierarchical Analysis											
SIP	β	Order 1	β	ΔR^2	Order 2	β	ΔR^2	Order 3	β	ΔR^2	Order 4	β	ΔR^2
Mastery experience (ME)	0.51***	ME	0.68***	0.47	VE	0.41***	0.17	VP	0.49***	0.24	AS	0.59***	0.35
Vicarious experience (VE)	0.02	VE	0.18***	0.03	VP	0.28***	0.08	AS	0.41***	0.17	ME	0.37***	0.14
Verbal persuasion (VP)	0.24***	VP	0.18***	0.03	AS	0.41***	0.17	ME	0.35***	0.13	VE	0.16***	0.03
Affective states (AS)	0.11	**AS**	**0.07**	**0.01**	**ME**	**0.35*** **	**0.12**	**VE**	**0.02**	**0.00**	**VP**	**0.16*** **	**0.02**
Total R^2	0.54			0.54			0.54			0.54			0.54

Notes: * $p < .05$., ** $p < .01$., *** $p < .001$. SIP = simultaneous inclusion of predictors in the regression model; $\Delta R2$ = incremental proportion of variance described in GTSE. The variables inserted lastly into the models and making a unique contribution are shown in bold.

4. Discussion

The first research question sought to determine whether the two scales used in this study, the TEIP and the STSE, measure the same constructs in both Japan and Finland. The scalar invariance for the first-order factor model and the metric invariance for the second-order factor model were achived using MGCFA. These results confirm that the construct validity of the two scales is invariant across the two countries. A further important finding regarding scale validity was that the MTMM analysis confirmed that the newly developed STSE scale performed well psychometrically. As mentioned in the literature review, few existing scales meet this standard or measure all four sources of TSE (Morris et al., 2017), and our findings confirm the utility of this new tool for measuring and analyzing the sources of TSE in future research.

The second research question sought to identify how the sources of self-efficacy affect TSE in both countries. This is the first study to use the Cholesky decomposition approach to explore the independent effects of these sources on TSE. The analysis revealed medium to high correlations between the source factors, indicating that the four sources overlap or mediate each other. This finding aligns with Bruce and Ross's (2008) finding that TSE is affected by the sources in combination. Because the four sources are themselves highly correlated, the issue of multicollinearity arose when conducting a multiple regression analysis in the SEM. However, the approach adopted enabled us to address this issue and to identify the unique contribution of each source.

Hierarchical regression models using the Cholesky decomposition revealed that ME was an independent source and the most powerful in relation to TSE for inclusive practices in both Japan and Finland. This is consistent not only with Hypothesis 1 but also with Bandura (1997) and other previous studies (Bruce & Ross, 2008; Milner, 2002).

The second source that made a unique contribution to GTSE in both countries was VP, although the effect was negative in Japan. These results are in agreement with Milner's (2002) findings, which showed that verbal feedback from students, parents, and colleagues was indispensable for TSE. This significant effect of VP in Finland appears to contradict the TALIS results, in which 91.9% of Finnish teachers reported that they had never had formal appraisal by

other teachers, and fewer teachers than the OECD average reported having received feedback by the following methods: (a) classroom observation (46.2%); (b) student surveys (26.2%); (c) assessment of teacher's content knowledge (25.9%); (d) analysis of student test scores (27.6%); (e) self-assessment of teacher's work (20.8%); and (f) surveys or discussion with parents (37.4%) (OECD, 2014). However, the TALIS study is limited in that it asks only about the above types of feedback, which might take other forms. In Finland, for example, teachers receive feedback through "individual developmental dialogue" with school leaders (OECD, 2014). Correspondingly, as there are no nationally regulated frameworks for teacher evaluation in Finland, teachers may receive informal feedback from colleagues rather than formal appraisal (OECD, 2014). What we wish to underscore here is that while these measures reflect the perceived influence of verbal persuasion on TSE, this does not mean that more verbal persuasion would necessarily lead to higher TSE. Rather, as mentioned earlier, the effectiveness of verbal persuasion depends both on who delivers it and how it is delivered. Engelbrecht, Savolainen, Nel, Koskela, and Okkolin (2017) contend that Finnish schools have a "well-developed learning support network," where teachers can receive daily (and mostly positive) feedback from colleagues, as well as from teachers specifically trained in special needs education. Further work is required to explore how such messages are framed and what kind of relationship exists between teachers and other staff in Finnish schools.

In contrast, the results suggest that VP had a negative effect on GTSE in Japan. This may be explained in part by the fact that Japanese teachers received feedback from principals (75.2%) and from the school management team (64.5%) more often than the OECD average (principals: 54.3%, school management team: 49.3%) (OECD, 2014). In a related vein, Tokyo Metropolitan School Personnel in Service Training Center (2007) conducted a survey of novice teachers and reported that only about 35% of those working in primary schools considered advice from principals and school management team to be helpful for problem solving and self development. As mentioned above, the role of verbal persuasion in enhancing TSE depends crucially on the relationship between group members and how the message is delivered (Morris et al., 2017). In light of Japan's hierarchical society (Nishimura, Nevgi, & Tella, 2008) and its teacher evaluation system (MEXT, 2014), it seems probable that verbal persuasion delivered by a principal or by a member of the school management team is seen as a formal appraisal for the purpose of teacher evaluation rather than as positive feedback to improve classroom teaching, especially among younger teachers. This does not

mean, however, that Japanese teachers receive no positive feedback at school. For example, there is evidence that about 80% of the novice teachers found advice from colleagues and mentors helpful when they encountered difficulties. This confirms the importance of how persuasive messages are framed; as some researchers have suggested, opportunities to receive positive and constructive feedback based on "collegiality" (Little, 1982) may be essential for Japanese teachers (Goto, 2014; Tsukiyama, 2006).

Although the correlation of VE and AS with GTSE was small for the Japanese data and medium to high for the Finnish data, VE and AS made no unique contribution to GTSE, and Hypothesis 2 was not supported. These results may indicate that the two sources (vicarious experience and psychological and affective states) do not independently or directly predict TSE for inclusive practices but instead mediate or moderate the other sources that affect TSE. In relation to the independent contribution of vicarious experience to self-efficacy, previous studies have reported inconsistent results. While some researchers have argued for a predictable relationship between vicarious experience and self-efficacy (Klassen, 2004b; Matsui, Matsui, & Ohnishi, 1990), others have found no such relationship (Anderson & Betz, 2001; Gainor & Lent, 1998). One possible explanation for this inconsistency may be that the effect of vicarious experience on self-efficacy is highly dependent on contextual factors such as the characteristics and relationships of group members (Bandura, 1997; Usher & Pajares, 2008). In the present context, the result may reflect (a) the difficulty of finding a role model in their own school with similar ability and personal attributes or (b) the lesser impact of vicarious experience on TSE at this developmental stage, where teachers have already established their own professional identity. Another contextual explanation could be related to the findings in the previous studies that vicarious experience may be more influential when the task is novel and the achievement goal is uncertain (Bandura, 1997; Chen & Usher, 2013) and/or during transitional periods (e.g., when children are moving from primary to lower secondary school) (Eccles, Midgley, & Adler, 1984; Usher & Pajares, 2008). For this reason, it seems possible that vicarious experience have some impact on TSE for inclusive practices because the concept of inclusive education is relatively new for teachers in both countries, and they are in a period of transition from separate special education to inclusive education. On that basis, another possible explanation, echoing Morris et al. (2017), is that teachers have insufficient opportunities to observe their colleagues, so limiting the influence of vicarious experiences on TSE. This view finds support

from the OECD (2014) TALIS study, which reported that only 5.1% of Finnish teachers and 29.8% of Japanese teachers indicated having participated in mentoring and peer observation in the previous 12 months. One means of increasing teachers' daily opportunities to observe and learn from their school colleagues is to implement co-teaching (Roth, Masciotra, & Boyd, 1999). Although this practice is increasing in Finland, the method is more frequently used by resource room and special class teachers (Saloviita & Takala, 2010). Among general education teachers, co-teaching still appears to be used less frequently (Saloviita & Takala, 2010). In contrast, the concept of "co-teaching" is not widely known in Japan, although the term "team teaching" is often used and implemented, and people sometimes use these terms interchangeably (Yamasaki, 2013). Further research should investigate how often teachers observe other teachers and how vicarious experience affects TSE for inclusive practices.

In relation to affective and psychological states, our results align with Poulou (2007), who showed that affective state is not itself a predictor of self-efficacy but rather mediates self-efficacy through cognitive processes. In the present study, this can perhaps be explained as a methodological problem; the items related to affective states required participants to indicate the extent to which "the feelings teaching has aroused" affected their beliefs about their teaching ability in each domain, which may seem too vague a question, especially as the influence of affective and psychological states is not episodic but ongoing (Morris et al., 2017). For that reason, future studies should ask more directly about specific states and how they affect particular aspects of TSE.

A final important finding is that the four sources explained only 15% of the variance of GTSE in the Japanese sample but explained 54% in the Finnish sample. This suggests that other sources of TSE for inclusive practices may exert a more powerful influence in Japan. There are some variables which could be other sources of TSE based on previous studies. For instance, Morris et al. (2017) demonstrated that teachers' content and pedagodical knowledge can improve their sense of self-efficacy, even though there is ongoing discussion whether the mastery of knowledge is an original source of TSE or derived from the identified four sources of self-efficacy. It has been suggested that the Japanese teachers reported considerable anxiety about implementing inclusive education because of the lack of knowledge and skills (Ueno & Nakamura, 2011). Thus, the mastery of knowledge regarding inclusive education might add some variation to Japanese teachers' self-

efficacy for inclusive practices. Another possible variable is a sense of collective efficacy in the school where teachers work. Goddard and Goddard (2001) found a significant positive relationship between TSE and collective efficacy and indicated that social influence shapes TSE considering Bandura's (1997) social cognitive theory. As previously noted, group harmony and "we" consciousness are highly important in collectivist culture (Hofstede, 2001; Markus & Kitayama, 1991), it is most probable that collective efficacy has stronger impact on TSE in Japanese context. Furthermore, in an open-ended questionnaire-based survey of teachers in Shanghai, Cheung (2008) found that students' and parents' confidence or respect was one of the most commonly cited factors contributing to TSE. This may be also the case in Japan because, in general, *shinyo* ("trustworthiness") is central to Japanese social morality (Lebra, 1976), and *sonkei* ("respect for others") is positively associated with affirmation of self-other relationships (Markus & Kitayama, 1991). As the confidence or respect of others can be broadly regarded as verbal persuasion, further research is needed on sources of TSE in non-Western contexts.

5. Limitations and future research

The aims of the present study were to test the construct validity of two scales and to examine the unique contribution of each of the four sources of TSE for inclusive practices. The findings reported here shed new light on how to measure and analyze the sources of TSE. In addition, although the results did not support Hypothesis 2, we were able to confirm that mastery experience is the most essential source of TSE for inclusive practices in both Japan and Finland, and that verbal persuasion may work differently in these differing ethnic contexts. While the findings contribute in a number of ways to the existing literature, the study has some limitations. The first of these is the generalizability of these results; for instance, both datasets were collected using convenience sampling, especially the Japanese sample, which included teachers from only one region. Moreover, the observed negative effect of VP on GTSE in the Japanese data should be interpreted with caution, as there remains a possibility that this result was a matter of chance. Further studies involving more samples from the same population are needed to assess the generalizability of these results. Similarly, differences between the sample sizes and data collection periods in Japan and Finland may adversely affect the comparability of these data, and future research should be designed to

gather a similar volume of data at the same time point.

A second limitation relates to the nature of the STSE scale. For example, the VE items in the present study asked participants to rate the extent to which "observations on other teachers having done well" affected their own abilities in the different teaching domains. However, according to Bandura (1997, pp. 93–95), "symbolic modeling" and "self-modeling" that utilizes recent technologies may be a source of vicarious information, as asking such questions may confine participants' reports to specific types of experience (Morris et al., 2017). In addition, the VE items did not ask whom they observed, what characteristics the model had and what kind of relationships the participant had with the model. As previously noted, those contextual factors are highly related to effect of vicarious experience (Bandura, 1997; Usher & Pajares, 2008). Thus, further studies, which take these variables into account, will need to be undertaken to find what characteristics are most relevant to teachers and how they capture the information about the characteristics in different countries. Similarly, the VP items used here did not specify from whom (e.g., colleagues, students, parents) comments were provided, and the AS items did not ask whether their feelings were positive or negative. Overall, despite confirmation of its sound psychometrics, the STSE scale may need to be modified in order to more accurately measure TSE sources and their various facets.

A third limitation of this study was that the mediating and moderating effects of the four sources on TSE were not addressed. Although our findings support previous research indicating that the four sources affect TSE in combination rather than independently (Bandura, 1997; Bruce & Ross, 2008; Morris et al., 2017), we did not examine mediating and moderating effects because of the complexity of the SEM model. To develop a fuller picture of the sources of TSE, further investigation of their interaction would be worthwhile, perhaps using a longitudinal research design.

Finally, our findings revealed that other sources may influence TSE for inclusive practices, especially in the case of Japanese teachers. As other sources that might predict TSE remain underspecified, mixed methods research based on sequential exploratory design (Creswell & Plano Clark, 2007) may provide a deeper and more detailed understanding.

6. Practical implications and conclusion

Notwithstanding the above limitations, the present findings have several practical implications. First of all, as the study confirms that mastery experience is the most powerful source of TSE for inclusive practices in both Japan and Finland, both governments should organize in-service and pre-service teacher training programs that will enable teachers to broaden their mastery experience in certain teaching domains. In particular, pre-service teacher education programs should provide opportunities to gain mastery experience through teaching practice so that novice teachers can enter this demanding job with confidence in their ability to implement inclusive education provisions. In addition, although no evidence was found that VE makes any independent contribution, this may be because teachers have limited opportunities to observe relevant role models. TSE may therefore be enhanced by providing more opportunities for modeling others, as well as for symbolic modeling and self-modeling, in both pre-service and in-service training programs. Finally, our findings suggest that verbal persuasion may have either positive or negative effects on TSE for inclusive practices, depending on the school context. As Finland's school learning support networks seem to enable teachers to receive persuasive information in a positive way (Engelbrecht et al., 2017), it would be worthwhile to explore the nature of these learning networks and how they work in Finnish schools and to utilize these insights to improve school working environments in other countries.

In sum, the present study confirmed the construct validity of the two scales in both Japan and Finland as a prerequisite for meaningful comparison. The cross-cultural analysis revealed interesting similarities and differences in terms of how the four sources of self-efficacy contribute to TSE for inclusive practices based on cultural and historical background. The reciprocal relationship between self (internal personal factors and behaviors) and society (external environment) outlined in Bandura's social cognitive theory (1997) makes it necessary for self-efficacy researchers to take contextual factors into account. Cross-cultural studies therefore offer useful insights into both the sources of self-efficacy and the development of TSE for inclusive practices, which will influence teachers' behavior to implement inclusive education.

Acknowledgements

This work was supported by Finnish Ministry of Education and Culture, Finland [grant number 35/529/2012].

References

Anderson, S. L., & Betz, N. E. (2001). Sources of social self-efficacy expectations: Their measurement and relation to career development. *Journal of Vocational Behavior*, *58*(1), 98–117. http://dx.doi.org/10.1006/jvbe.2000.1753

Avramidis, E., & Norwich, B. (2002). Teachers' attitudes towards integration/inclusion: A review of the literature. *European Journal of Special Needs Education*, *17*(2), 129–147. http:// dx.doi.org/10.1080/08856250210129056

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, *84*(2), 191–215. http:// dx.doi.org/10.1037/0033-295X.84.2.191

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York, NY: Freeman.

Björn, P. M., Aro, M. T., Koponen, T. K., Fuchs, L. S., & Fuchs, D. H. (2016). The many faces of special education within RTI frameworks in the United States and Finland. *Learning Disability Quarterly*, *39*(1), 58–66. http:// dx.doi.org/10.1177/0731948715594787

Borovoy, A. (2008). Japan's hidden youths: Mainstreaming the emotionally distressed in Japan. *Culture, Medicine and Psychiatry*, *32*(4), 552–576. http://dx.doi.org/10.1007/s11013-008-9106-2

Britner, S. L., & Pajares, F. (2006). Sources of science self-efficacy beliefs of middle school students. *Journal for Research in Science Teaching*, *43*, 485–499. https://dx.doi.org/10.1002/tea.20131

Brouwers, A., & Tomic, W. (2001). The factorial validity of scores on the teacher interpersonal self-efficacy scale. *Educational and Psychological Measurement*, *61*(3), 433–445. https://dx.doi.org/10.1177/00131640121971301

Bruce, C. D., & Ross, J. A. (2008). A model for increasing reform implementation and teacher efficacy: Teacher peer coaching in grades 3 and 6 mathematics. *Canadian Journal of Education*, *31*(2), 346–370. http://dx.doi.org/10.2307/20466705

Byrne, B. M. (2013). *Structural equation modeling with M plus: Basic concepts, applications, and programming*. London: Routledge.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, *56*(2), 81–105. http://dx.doi.org/10.1037/h0046016

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, *14*(3), 464–504. http://dx.doi.org/10.1080/10705510701301834

Chen, F. F., Sousa, K. H., & West, S. G. (2005). Teacher's corner: Testing measurement invariance of second-order factor models. *Structural Equation Modeling: A Multidisciplinary Journal*, *12*(3), 471–492. http://dx.doi.org/10.1207/s15328007sem1203_7

Chen, J. A., & Usher, E. L. (2013). Profiles of the sources of science self-efficacy. *Learning and Individual Differences*, *24*, 11–21. http://dx.doi.org/10.1016/j.lindif.2012.11.002

Cheung, H. Y. (2008). Teacher efficacy: A comparative study of Hong Kong and Shanghai Primary in-service teachers. *The Australian Educational Researcher*, *35*(1), 103–123. http://dx.doi.org/10.1007/BF03216877

Chiu, M. M., & Klassen, R. M. (2009). Calibration of reading self-concept and reading achievement among 15-year-olds: Cultural differences in 34 countries. *Learning and Individual Differences*, *19*(3), 372–386. http://dx.doi.org/10.1016/j.lindif.2008.10.004

Creswell, J. W., & Plano Clark, V. L. (2007). *Designing and conducting mixed methods research.* Thousand Oaks, CA: Sage Publications.

de Boer, A., Jan Pijl, S., & Minnaert, A. (2011). Regular primary schoolteachers' attitudes towards inclusive education: A review of the literature. *International Journal of Inclusive Education*, *15*(3), 331–353. http://dx.doi.org/10.1080/13603110903030089

de Jong, P. F. (1999). Hierarchical regression analysis in structural equation modeling. *Structural Equation Modeling*, *6*(2), 198–211. http://dx.doi.org/10.1080/10705519909540128

Eccles, J. S., Midgley, C., & Adler, T. (1984). Grade-related changes in the school environment: Effects on achievement motivation. In J. Nicholls (Ed.), *Advances in motivation and achievement: The development of achievement motivation* (pp. 238–331). Greenwich, CT: JAI Press.

Engelbrecht, P., Savolainen, H., Nel, M., Koskela, T., & Okkolin, M. A. (2017). Making meaning of inclusive education: classroom practices in Finnish and South African classrooms. *Compare: A Journal of Comparative and International Education*, *47*(5), 684-702. http://dx.doi.org/10.1080/03057925.2016.1266927

FNBE [Finnish National Board of Education]. (2016). *National Core Curriculum for Basic Education 2014*. Helsinki: Finnish National Board of Education.

Forlin, C., Kawai, N., & Higuchi, S. (2015). Educational reform in Japan towards inclusion: Are we training teachers for success? *International Journal of Inclusive Education*, *19*(3), 314–331. http://dx.doi.org/10.1080/13603116.2014.930519

Forlin, C., Sharma, U., & Loreman, T. (2014). Predictors of improved teaching efficacy following basic training for inclusion in Hong Kong. *International Journal of Inclusive Education*, *18*(7), 718-730. https://dx.doi.org/10.1080/13603116.2013.819941

Fujii, N. (2014). Inkuru-shibukyoikushisutemukochiku no hokosei ni kansuru kento: Kyoshokuin ni taisuru ki-wa-do no ninchidochosa wo toshite [Study into the directions in building the inclusive education system: Through teaching staff survey on keyword awareness]. *Bulletin of Center for Educational Research and Practice, Akita University*, *36*, 89–98.

Gainor, K. A., & Lent, R. W. (1998). Social cognitive expectations and racial identity attitudes in predicting the math choice intentions of black college students. *Journal of Counseling Psychology*, *45*(4), 403–413. http://dx.doi.org/10.1037/0022-0167.45.4.403

Goddard, R., & Goddard, Y. (2001). A multilevel analysis of the relationship between teacher and collective efficacy in urban schools. *Teaching and Teacher Education*, *17*(7), 807–818. http://dx.doi.org/10.1016/S0742-051X(01)00032-4

Goto, I. (2014). *Shogakkoshoninkyoyu no seicho hattatsu o sasaeru atarashii ikuseiron* [*New training theory to support the growth and development of novice elementary school teachers*]. Tokyo: Gakujutsu Shuppankai.

Hakala, K., Björnsdóttir, K., Lappalainen, S., Jóhannesson, I. Á., & Teittinen, A. (2018). Nordic perspectives on disability studies in education: A review of research in Finland and Iceland. *Education Inquiry*, *9*(1), 1–19. http://dx.doi.org/10.1080/20004508.2017.1421390

Hall, E. T. (1976). *Beyond culture*. New York, NY: Anchor Books/Doubleday.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*(1), 81–112. http://dx.doi.org/10.3102/003465430298487

Hofstede, G. (2001). *Culture's consequences: Comparing values, behaviors, institutions and organizations across nations.* Thousand Oaks, CA: Sage Publications.

Horst, P., Wallin, P., Guttman, L., Wallin, F. B., Clausen, J. A., Reed, R., & Rosenthal, E. (1941). *The prediction of personal adjustment: A survey of logical problems and research techniques, with illustrative application to problems of vocational selection, school success, marriage, and crime.* New York, NY: Social Science Research Council.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. http://dx.doi.org/10.1080/10705519909540118

Ichikawa, N. (2016). Yomikakishogai no aru kodomo ni taisuru shien: Gakko ni okeru ”imasubeki” shien toha [Support to children with dyslexia: What is the effective support in school now?]. *Bulletin of Shiraume Gakuen University and College*, *52*, 73–90.

Jahnukainen, M. (2011). Different strategies, different outcomes? The history and trends of the inclusive and special education in Alberta (Canada) and in Finland. *Scandinavian Journal of Educational Research*, *55*(5), 489–502. http://dx.doi.org/10.1080/00313831.2010.537689

Jahnukainen, M. (2015). Inclusion, integration, or what? A comparative study of the school principals' perceptions of inclusive and special education in Finland and in Alberta, Canada. *Disability & Society*, *30*(1), 59–72. http://dx.doi.org/10.1080/09687599.2014.982788

Jahnukainen, M., & Itkonen, T. (2016). Tiered intervention: History and trends in Finland and the United States. *European Journal of Special Needs Education*, *31*(1), 140–150. http://dx.doi.org/10.1080/08856257.2015.1108042

Kagitçibasi, C. (1997). Individualism and collectivism. In J. W. Berry, M. H. Segall, & C. Kagitçibasi (Eds.), *Handbook of cross-cultural psychology, Volume 3: Social behavior and applications* (pp. 1–50). Needham Heights, MA: Allyn & Bacon.

Kivirauma, J., Klemelä, K., & Rinne, R. (2006). Segregation, integration, inclusion: The ideology and reality in Finland. *European Journal of Special Needs Education*, *21*(2), 117–133. http://dx.doi.org/10.1080/08856250600600729

Klassen, R. M. (2004a). A cross-cultural investigation of the efficacy beliefs of South Asian immigrant and Anglo Canadian nonimmigrant early adolescents. *Journal of Educational Psychology*, *96*(4), 731–742. http://dx.doi.org/10.1037/0022-0663.96.4.731

Klassen, R. M. (2004b). Optimism and realism: A review of self-efficacy from a cross-cultural perspective. *International Journal of Psychology*, *39*(3), 205–230. http://dx.doi.org/10.1080/00207590344000330

Klassen, R. M., Bong, M., Usher, E. L., Chong, W. H., Huan, V. S., Wong, I. Y. F., & Georgiou, T. (2009). Exploring the validity of a teachers' self-efficacy scale in five countries. *Contemporary Educational Psychology*, *34*(1), 67–76. http://dx.doi.org/10.1016/j.cedpsych.2008.08.001

Klassen, R. M., Tze, V. M. C., Betts, S. M., & Gordon, K. A. (2011). Teacher efficacy research 1998-2009: Signs of progress or unfulfilled promise? *Educational Psychology Review*, *23*(1), 21–43. http://dx.doi.org/10.1007/s10648-010-9141-8

Lebra, T. S. (1976). *Japanese patterns of behaviour*. Honolulu, HI: University of Hawaii Press.

Little, J. W. (1982). Norms of collegiality and experimentation: Workplace conditions of school success. *American Educational Research Journal*, *19*(3), 325–340. https://dx.doi.org/10.3102/00028312019003325

Lubin, A. (1957). Some formulae for use with suppressor variables. *Educational and Psychological Measurement*, *17*(2), 286–296. https://dx.doi.org/10.1177/001316445701700209

MacCallum, R. C., Browne, M. W., & Cai, L. (2006). Testing differences between nested covariance structure models: Power analysis and null hypotheses. *Psychological Methods*, *11*(1), 19–35. http://dx.doi.org/10.1037/1082-989X.11.1.19

MacKinnon, D. P., Krull, J. L., & Lockwood, C. M. (2000). Equivalence of the mediation, confounding and suppression effect. *Prevention Science*, *1*(4), 173–181. https://dx.doi.org/10.1023/A:1026595011371

Malinen, O.-P. (2014). *Improving teachers' sense of professional efficacy: Empirical investigation on the sources of teacher efficacy.* Paper presented at the meeting of the 2nd Global Teacher Education Summit, Beijing Normal University, China.

Malinen, O.-P., Savolainen, H., & Xu, J. (2013). Dimensions of teacher self-efficacy for inclusive practices among mainland Chinese pre-service teachers. *Journal of International Special Needs Education*, *16*(2), 82–93. http://dx.doi.org/10.9782/2159-4341-16.2.82

Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review*, *98*(2), 224–253. http://dx.doi.org/10.1037/0033-295X.98.2.224

Matsui, T., Matsui, K., & Ohnishi, R. (1990). Mechanisms underlying math self-efficacy learning of college students. *Journal of Vocational Behavior*, *37*(2), 225–238. http://dx.doi.org/10.1016/0001-8791(90)90042-Z

Meijer, C. J. W., & Foster, S. F. (1988). The effect of teacher self-efficacy on referral chance. *The Journal of Special Education*, *22*(3), 378–385. http://dx.doi.org/10.1177/002246698802200309

MEXT. (2012). *Kyoseishakai no keisei ni muketa inkurushibukyoikushisutemukochiku no tameno tokubetsushienkyoiku no suishin (Hokoku)* [*Promoting special needs education to construct inclusive education system that formulates convivial society (Report)*]. Retrieved on 16 May, 2018 from: http://www.mext.go.jp/b_menu/shingi/chukyo/chukyo3/044/attach/1321669.htm

MEXT. (2014). *Kyoshokuinhyoka no genjo* [*The current situation of teacher evaluation*]. Retrieved on 18 June, 2018 from: http://www.mext.go.jp/b_menu/shingi/chukyo/chukyo3/052/siryo/__icsFiles/afieldfile/2016/01/05/1365651_10.pdf

MEXT. (2017). *Gakkokyointokeichosa: Heisei 28 nendokekka no gaiyo* [*Statistics in school teachers: Outline of the survey results in 2016*]. Retrieved on 12 March, 2018 from: http://www.mext.go.jp/b_menu/toukei/chousa01/kyouin/kekka/k_detail/1395309.htm

Mills, N. (2011). Teaching assistants' self-efficacy in teaching literature: Sources, personal assessments, and consequences. *Modern Language Journal*, *95*(1), 61–80. http://dx.doi.org/10.1111/j.1540-4781.2010.01145.x

Milner, H. R. (2002). A case study of an experienced English teacher's self-efficacy and persistence through "crisis" situations: Theoretical and practical considerations. *The High School Journal*, *86*(1), 28–35.

Milner, H. R., & Hoy, A. W. (2003). A case study of an African American teacher's self-efficacy, stereotype threat, and persistence. *Teaching and Teacher Education*, *19*(2), 263–276. http://dx.doi.org/10.1016/S0742-051X(02)00099-9

Ministry of Education of Finland. 2007. *Special education strategy*. *Reports of the Ministry of Education, Finland, no. 47*. Helsinki, Finland: Ministry of Education of Finland.

Mitchell, D. (2005). Introduction: Sixteen propositions on the contexts of inclusive education. In D. Mitchell (Ed.), *Contextualizing inclusive education: Evaluating old and new international perspectives* (pp. 1–21). London: Routledge.

Morris, D. B., Usher, E. L., & Chen, J. A. (2017). Reconceptualizing the sources of teaching self-efficacy: A critical review of emerging literature. *Educational Psychology Review*, *29*(4), 795–833. http://dx.doi.org/10.1007/s10648-016-9378-y

Muthén, L. K., & Muthén, B. O. (2012). Mplus version 7 [Statistical software]. Los Angeles, CA: Muthén & Muthén.

Nishimura, S., Nevgi, A., & Tella, S. (2008). Communication style and cultural features in high/low context communication cultures: A case study of Finland, Japan and India. In A. Kallioniemi

(Ed.), *Uudistuva ja kehittyvä ainedidaktiikka: Ainedidaktinen symposiumi 8.2.2008 Helsingissä* (pp. 783–796). Helsinki: Helsingin yliopisto. Retrieved on 10 July, 2018 from: http://www.helsinki.fi/~tella/nishimuranevgitella299.pdf

OECD [The Organization for Economic Cooperation and Development]. (2013). *Key findings from the Teaching and Learning International Survey (TALIS)*. Retrieved on 9 July, 2018 from: http://www.oecd.org/finland/TALIS-2013-country-note-Finland.pdf

OECD. (2014). *TALIS 2013 results: An international perspective on teaching and learning*. Paris: OECD Publishing. http://dx.doi.org/10.1787/9789264196261-en

OECD. (2018). *Foreign population (indicator)*. Retrieved on 9 July, 2018 from: http://dx.doi.org/10.1787/16a914e3-en

Ogiso, H., & Tsuzuki, S. (2016). Tokubetsushienkyoikutaisei to inkurushibukyoikushisutemu ni kansuru ichikosatsu [Consideration on the special support education system and inclusive education system]. *Shogaishakyoiku・fukushigakukenkyu*, *12*, 155–163.

Pajares, F., Johnson, M. J., & Usher, E. L. (2007). Sources of writing self-efficacy beliefs of elementary, middle, and high school students. *Research in the Teaching of English*, *42*, 104–120.

Palmer, D. (2011). Sources of efficacy information in an inservice program for elementary teachers. *Science Education*, *95*(4), 577–600. http://dx.doi.org/10.1002/sce.20434

Pesonen, H., Itkonen, T., Jahnukainen, M., Kontu, E., Kokko, T., Ojala, T., & Pirttimaa, R. (2015). The implementation of new special education legislation in Finland. *Educational Policy*, *29*(1), 162–178. http://dx.doi.org/10.1177/0895904814556754

Poulou, M. (2007). Personal teaching efficacy and its sources: Student teachers' perceptions. *Educational Psychology*, *27*(2), 191–218. http://dx.doi.org/10.1080/01443410601066693

Ross, J., & Bruce, C. (2007). Professional development effects on teacher efficacy: Results of randomized field trial. *Journal of Educational Research*, *101*(1), 50–60. http://dx.doi.org/10.3200/JOER.101.1.50-60

Roth, W. M., Masciotra, D., & Boyd, N. (1999). Becoming-in-the-classroom: A case study of teacher development through coteaching. *Teaching and Teacher Education*, *15*(7), 771–784. https://dx.doi.org/10.1016/S0742-051X(99)00027-X

Savolainen, H., Engelbrecht, P., Nel, M., & Malinen, O.-P. (2012). Understanding teachers' attitudes and self-efficacy in inclusive education: Implications for pre-service and in-service teacher education. *European Journal of Special Needs Education*, *27*(1), 51–68. http://dx.doi.org/10.1080/08856257.2011.613603

Saloviita, T., & Takala, M. (2010). Frequency of co-teaching in different teacher categories. *European Journal of Special Needs Education*, *25*(4), 389-396. https://dx.doi.org/10.1080/08856257.2010.513546

Schunk, D. H. (1984). Sequential attributional feedback and children's achievement behaviors. *Journal of Educational Psychology*, *76*(6), 1159–1169. http://dx.doi.org/10.1037/0022-0663.76.6.1159

Schwartz, S. H. (1999). A theory of cultural values and some implications for work. *Applied Psychology*, *48*(1), 23–47. https://dx.doi.org/10.1111/j.1464-0597.1999.tb00047.x

Sharma, U., Aiello, P., Pace, E. M., Round, P., & Subban, P. (2018). In-service teachers' attitudes, concerns, efficacy and intentions to teach in inclusive classrooms: an international comparison of Australian and Italian teachers. *European Journal of Special Needs Education*, *33*(3), 437-446. https://dx.doi.org/10.1080/08856257.2017.1361139

Sharma, U., Loreman, T., & Forlin, C. (2012). Measuring teacher efficacy to implement inclusive practices. *Journal of Research in Special Educational Needs*, *12*(1), 12–21. http://dx.doi.org/10.1111/j.1471-3802.2011.01200.x

Smith, P. B., Bond, M. H., & Kâğıtçıbaşı, Ç. (2006). *Understanding social psychology across cultures: Living and working in a changing world*. London: SAGE.

Soodak, L. C., & Podell, D. M. (1993). Teacher efficacy and student problem as factors in special education referral. *Journal of Special Education*, *27*(1), 66–81. http://dx.doi.org/10.1177/002246699302700105

Spielberger, C. (Ed.). (2004). *Encyclopedia of applied psychology*. San Diego, CA: Elsevier Academic Press.

Stevens, T., Olivárez, Jr. A., & Hamman, D. (2006). The role of cognition, motivation, and emotion in explaining the mathematics achievement gap between Hispanic and White students. *Hispanic Journal of Behavioral Sciences*, *28*(2), 161–186. https://dx.doi.org/10.1177/0739986305286103

Tokyo Metropolitan School Personnel in Service Training Center. (2007). *Shoninshakyoyuikusei ni kansuru shidoshiryo* [*Instruction guidelines for beginning teacher training*]. Retrieved on 25

May, 2018 from: http://www.kyoiku-kensyu.metro.tokyo.jp/09seika/reports/files/bulletin/h18/h18_proj01_1.pdf

Triandis, H. C. (2001). Individualism and collectivism: Past, present, and future. In D. Matsumoto (Ed.), *The handbook of culture and psychology* (pp. 35–50). New York, NY: Oxford University Press.

Tschannen-Moran, M., & Hoy, A. W. (2007). The differential antecedents of self-efficacy beliefs of novice and experienced teachers. *Teaching and Teacher Education*, *23*(6), 944–956. http://dx.doi.org/10.1016/j.tate.2006.05.003

Tschannen-Moran, M., Hoy, A. W., & Hoy, W. K. (1998). Teacher efficacy: Its meaning and measure. *Review of Educational Research*, *68*(2), 202–248. https://dx.doi.org/10.3102/00346543068002202

Tsukiyama, T. (2006). Ima o ikiru seinenkyoshi o sasaerumono: Tomoni seicho suru yorokobi [Supporting novice teachers who are living in a present society]. *Hiroba, Quarterly Journal of Education Center of Kyoto*, (147). Retrieved on 10 July, 2018 from: http://www.kyoto-kyoiku.com/hiroba2/hiroba147/hiroba147tukiyama.htm

Ueno, K., & Nakamura, K. (2011). Inkuru-jonkyoiku ni taisuru tsujogakkyukyoin no ishiki ni tsuite [A study of awareness of "Inclusion Education" among regular-class teachers in elementary schools]. *Journal of Health and Sports Science Juntendo University*, *3*(2), 112–117.

UNESCO [The United Nations Educational, Scientific and Cultural Organization]. (1994). *The Salamanca statement and framework for action on special needs education.* Retrieved on 29 October, 2017 from: http://www.unesco.org/education/pdf/SALAMA_E.PDF

United Nations. (2006). *Convention on the rights of persons with disabilities.* Retrieved on 29 October, 2017 from: http://www.un.org/disabilities/documents/convention/convoptprot-e.pdf

United Nations General Assembly. (2015). *Transforming our world: The 2030 agenda for sustainable development*. Retrieved on 25 May, 2017 from: http://www.un.org/ga/search/view_doc.asp?symbol=A/70/L.1

Usher, E. L., & Pajares, F. (2008). Sources of self-efficacy in school: Critical review of the literature and future directions. *Review of Educational Research*, *78*(4), 751–796. http://dx.doi.org/10.3102/0034654308321456

Vieluf, S., Kunter, M., & van de Vijver, F. J. (2013). Teacher self-efficacy in cross-national perspective. *Teaching and Teacher Education*, *35*, 92–103. https://dx.doi.org/10.1016/j.tate.2013.05.006

Wheatley, K. F. (2005). The case for reconceptualizing teacher efficacy research. *Teaching and Teacher Education*, *21*(7), 747–766. http://dx.doi.org/10.1016/j.tate.2005.05.009

Wyatt, M. (2014). Towards a re-conceptualization of teachers' self-efficacy beliefs: Tackling enduring problems with the quantitative research and moving on. *International Journal of Research and Method in Education*, *37*(2), 166–189. http://dx.doi.org/10.1080/1743727X.2012.742050

Yada, A., Tolvanen, A., & Savolainen, H. (2018). Teachers' attitudes and self-efficacy on implementing inclusive education in Japan and Finland: A comparative study using multi-group structural equation modelling. *Teaching and Teacher Education*, *75*, 343-355. https://dx.doi.org/10.1016/j.tate.2018.07.011

Yada, A., & Savolainen, H. (2017). Japanese in-service teachers' attitudes towards inclusive education and self-efficacy for inclusive practices. *Teaching and Teacher Education*, *64*, 222–229. http://dx.doi.org/10.1016/j.tate.2017.02.005

Yamasaki, T. (2013). Kyoshi no jissentekishidoryoku no keisei ni okeru koteichingu no kanosei [The possibility of co-teaching for developing teachers' instructional skills in practices]. *Journal of Elementary Education and Curriculum*, *1*, 21–31.

Appendix

The Sources of Teacher Self-Efficacy (STSE) scale

Evaluate how much the following factors have affected your view of your own abilities on four different aspects of work as a teacher.

Instructions: You can select any one of nine alternatives, which range between (1) "Not at all" and (9) "Very much". Option (5) "To some extent" represents the middle point between the two extremes.

1. Teaching learning contents (e.g. ability to plan learning assignments that are challenging enough for students, ability to assess students' understanding). How much have the following affected your view on these abilities:

		Not at all		Slightly		To some extent		Moderately		Very much
1.1.	My own experiences on how well I have succeeded/done	1	2	3	4	5	6	7	8	9
1.2.	My observations on other teachers having done well	1	2	3	4	5	6	7	8	9
1.3.	Comments on my work that I have received from other people	1	2	3	4	5	6	7	8	9
1.4.	The feelings teaching has aroused	1	2	3	4	5	6	7	8	9

2. Classroom management and behavior management of individual students (e.g. ability to calm and prevent disruptive behaviors, ability to get students to follow classroom rules). How much have the following affected your view on these abilities:

		Not at all		Slightly		To some extent		Moderately		Very much
2.1.	My own experiences on how well I have succeeded/done	1	2	3	4	5	6	7	8	9
2.2.	My observations on other teachers having done well	1	2	3	4	5	6	7	8	9
2.3.	Comments on my work that I have received from other people	1	2	3	4	5	6	7	8	9
2.4.	The feelings teaching has aroused	1	2	3	4	5	6	7	8	9

3. Collaboration (e.g. ability to collaborate with families of students, ability to with other professionals in the school, ability to work with professionals outside of school). How much have the following affected your view on these abilities:

		Not at all		Slightly		To some extent		Moderately		Very much
3.1.	My own experiences on how well I have succeeded/done	1	2	3	4	5	6	7	8	9
3.2.	My observations on other teachers having done well	1	2	3	4	5	6	7	8	9
3.3.	Comments on my work that I have received from other people	1	2	3	4	5	6	7	8	9
3.4.	The feelings teaching has aroused	1	2	3	4	5	6	7	8	9

4. Supporting students' school motivation (e.g. ability to motivate students who show little interest in school work, ability to support students beliefs in their own abilities). How much have the following affected your view on these abilities:

		Not at all		Slightly		To some extent		Moderately		Very much
4.1.	My own experiences on how well I have succeeded/done	1	2	3	4	5	6	7	8	9
4.2.	My observations on other teachers having done well	1	2	3	4	5	6	7	8	9
4.3.	Comments on my work that I have received from other people	1	2	3	4	5	6	7	8	9
4.4.	The feelings teaching has aroused	1	2	3	4	5	6	7	8	9