Exploring the validity of a teachers' self-efficacy scale in five countries

Robert M. Klassen a,* , Mimi Bong b , Ellen L. Usher c , Wan Har Chong d , Vivien S. Huan d , Isabella Y.F. Wong d , Tasos Georgiou e

a University of Alberta, Department of Educational Psychology, 6-102 Education North, Edmonton, AB, Canada T6G 2C5
b Department of Education, Korea University, Anam-dong, Seongbuk-gu, Seoul 136-701, Korea
c Educational and Counseling Psychology, University of Kentucky, 249 Dickey Hall, Lexington, KY 40506-0017, USA
d Psychological Studies Academic Group, National Institute of Education, 1 Nanyang Walk, Singapore 637616
e Corner Thoucydides and Kimon, 1434 Nicosia, Cyprus

Abstract

The purpose of this article was twofold. The first purpose was to test the validity of the Teachers’ Sense of Self-Efficacy Scale (TSES) in five settings—Canada, Cyprus, Korea, Singapore, and the United States. The second purpose was, by extension, to establish the importance of the teacher self-efficacy construct across diverse teaching conditions. Multi-group confirmatory factor analysis was used to better understand the measurement invariance of the scale across countries, after which the relationship between the TSES, its three factors, and job satisfaction was explored. The TSES showed convincing evidence of reliability and measurement invariance across the five countries, and the relationship between the TSES and job satisfaction was similar across settings. The study provides general evidence that teachers’ self-efficacy is a valid construct across culturally diverse settings and specific evidence that teachers’ self-efficacy showed a similar relationship with teachers’ job satisfaction in five contrasting settings.

1. Introduction

Teachers across the world face a host of new and ongoing demands resulting from increasing workloads, shifting policies and expectations, and societal changes. The beliefs teachers hold about their capabilities to face these challenges play a strong role in influencing student learning and teachers’ job commitment (Tschannen-Moran & Woolfolk Hoy, 2001). In recent years, teachers’ self-efficacy (TSE)—beliefs teachers hold that they can positively influence student learning—has been shown to demonstrate a profound influence on the daily lives of teachers and their students. Although considerable research has shown that TSE has a strong influence on teachers and students, few studies have explored the validity of TSE across groups of teachers from different settings. In particular, international comparisons provide researchers with new knowledge about the universality and generalizability of important psychological constructs, and allow future investigations to include the newly validated constructs in a more diverse range of settings (e.g., Marsh & Hau, 2004). The purpose of the present investigation is to test the validity of a TSE measure across teaching levels (elementary/middle school and secondary) in five countries representing three geographically and culturally contrasting regions: North America (Canada and the United States), East Asia (Korea and Singapore) and Europe (Cyprus).

1.1. Teachers’ self-efficacy

According to Bandura’s (1986) social cognitive theory, self-efficacy beliefs refer to individuals’ beliefs about their capabilities to successfully carry out a particular course of action. Considerable research supports the claim that self-efficacy is an important influence on human achievement in a wide variety of settings, including education, health, sports, and work (Bandura, 1997). In educational contexts, research has shown that students’ self-efficacy plays an important role in influencing achievement and behavior, but there is increasing evidence that teachers’ sense of self-efficacy also plays a key role in influencing important academic outcomes. Teachers’ self-efficacy is related to higher levels of student achievement and student motivation, and has been shown to influence teachers’ instructional practices, enthusiasm, commitment, and teaching behavior (Skaoilvk & Skoailvk, 2007; Tschannen-Moran & Woolfolk Hoy, 2001; Wolters & Daugherty, 2007). Teachers with low levels of self-efficacy experience more difficulties with student
misbehavior, are pessimistic about student learning, and experience higher levels of job-related stress and lower levels of job satisfaction (Bandura, 1997; Caprara, Barbaranelli, Borgogni, & Steca, 2003; Caprara, Barbaranelli, Steca, & Malone, 2006; Lee, Dedrick, & Smith, 1991). The relationship between TSE and job satisfaction is especially important because job satisfaction has been shown to be significantly related to job performance across a wide range of work settings (Judge, Thoresen, Bono, & Patton, 2001), and more importantly, is considered a critical element affecting teachers’ attitudes and efforts in their daily work with children (Caprara et al., 2003). Exploring the relationship between TSE and job satisfaction may have implications for teachers’ job performance, and by extension, the academic achievement of students.

Research on TSE dates back to educational studies carried out by the RAND organization in the mid-1970s, when two questionnaire items were created to investigate teachers’ beliefs in their ability to influence student achievement (Tschannen-Moran, et al., 2003). Exploring the relationship between TSE and job satisfaction may be a frequent area of focus for researchers in the field of educational psychology.

1.2. Teachers’ self-efficacy in international perspective

Cross-national comparisons are useful theory-builders because they provide researchers with a “valuable heuristic basis to test the external validity and generalizability of their measures, theories, and models” (Marsh & Hau, 2004, p. 59) and because these comparisons offer a way of exploring the universality of psychological constructs and measures (Triandis, 1996).

Comparison of teacher motivation across diverse settings is educational because teaching practices and conditions show considerable variation within and across countries, and variations in teaching environments and teaching practices may influence teachers’ beliefs about their roles and responsibilities (Ho & Hau, 2004). For example, teachers in East Asian settings may have very different day-to-day working experiences than teachers in North American settings due to differences in teacher preparation, induction, professional development, and expectations for student behavior and achievement (Preus, 2007; Yeom & Ginsburg, 2007). Teachers in European settings like Cyprus may work under a more highly centralized education system than most North American teachers, and may be motivated to enter teaching for reasons different than American teachers (Zemblys & Papanastasiou, 2004). Although TSE may be a universally valid factor underlying teaching and learning, there has been little research exploring whether TSE operates in similar ways in settings that provide different teaching and learning experiences.

Almost all TSE studies to date have been conducted in a single, usually culturally Western setting, with scant attention paid to how self-efficacy operates in diverse contexts. Gorrell and colleagues (Gorrell & Hwang, 1995; Lin, Gorrell, & Taylor, 2002) conducted a series of studies investigating TSE in Korea and Taiwan, but their studies focused on pre-service teachers, and did not include cross-national comparisons. Ho and Hau (2004) conducted what might be the only study designed to explore TSE using a cross-national comparison. Their study with groups of teachers from Hong Kong and Australia borrowed items from Gibson and Dembo’s (1984) measure and from other studies resulting in a TSE measure that included components of personal teaching efficacy and general teaching efficacy.

In spite of the value of the cross-national analysis present by Ho and Hau (2004), the self-efficacy instrument they used may not be well aligned with accepted theoretical conceptions of self-efficacy. The majority of items assessed ability judgments (e.g., “I have very effective classroom management skills,” and “I am good at counseling students”) that bear closer resemblance to self-concept than self-efficacy. In addition, items composing the fourth factor, labeled “External Influences,” bear even less resemblance to self-efficacy (e.g., “A teacher is limited in what he/she can achieve in student discipline because of the current values of society”) (Ho & Hau, 2004, p. 317) because they do not assess teachers’ self-referent judgments, and more closely resemble the environmental obstacles included in the Gibson and Dembo (1984) GTE measure. Whereas self-concept reflects an assessment of skill and ability, and is formed through social and self-comparisons, self-efficacy is a judgment of capability to perform a particular task (Bong, 2006; Bong & Skaalvik, 2003; Pajares, 1996). Self-efficacy researchers have suggested that items designed to measure self-efficacy should be phrased “in terms of can do rather than will do. Can is a judgment of capability; will is a statement of intention” (Bandura, 2006, p. 308, italics in original). Self-efficacy instruments should focus not on assessing current skills, but on beliefs “about what one can do under different sets of conditions with whatever skills one possesses” (Bandura, 1997, p. 37). In light of the theoretical

---

1 We use the term international and cross-national rather than cross-cultural to describe the study because the primary defining comparison feature is country of residence, and cultural dimensions were not directly measured.
structures typically include a stronger focus on building teacher-student populations (e.g., K-7 and grades 6–8), (b) the two school systems often serve overlapping student relationships to a degree not typical in secondary schools (National Middle School Association, 2008), (c) the elementary and middle school samples shared similar demographic profiles and looked very similar on the TSE scales, and (d) combining the samples allowed for a greater range of analysis possibilities (e.g., CFA is a large sample technique in which sample size influences model quality; Gagné & Hancock, 2006). Comparisons using ANOVA showed no significant differences for mean age or teaching experience between elementary and middle school samples within Canada, Korea, or the United States. Comparisons of the levels of the four TSE scores (i.e., composite score and the three-factor scores) using ANOVA showed no significant differences in 11 of the 12 comparisons, with Canadian elementary teachers rating their TSE for student engagement (M = 27.94, SD = 3.91) higher than middle school teachers (M = 26.07, SD = 4.52), F(1, 202) = 8.57, p = .005. However, this difference can be considered non-significant when adjusted using a Bonferroni correction. Overall, the elementary and middle school teachers shared similar profiles in terms of age, teaching experience, and response patterns on the TSES.

The 709 elementary/middle school teachers were from four countries: Canada (n = 210), Cyprus (n = 210), Korea (n = 153), and the United States (n = 137). The Canadian elementary/middle school teachers were 76% female, with a mean age of 40.1 years (SD = 10.7), and taught in elementary schools (70%), middle schools (27%), and combined elementary-middle schools (3%), in 87% urban, and 13% suburban schools. The Cypriot teachers were 70% female, with a mean age of 33.2 years (SD = 9.0), and taught in 16 elementary schools described as 34% urban, 11% suburban, and 54% rural. Teachers from Korea were 78% female, were of Korean heritage (100%), had a mean age of 33.5 years (SD = 8.2), and taught in more than 40 schools, including elementary schools (18%) and middle schools (82%) in largely urban settings (97%). Teachers from the United States were volunteers from nine southeastern schools in a metropolitan area, and were 90% female, with a mean age of 38.9 years (SD = 11.5), and taught in elementary schools (55%), middle schools (43%), and combined elementary-middle schools (2%), in 21% urban, and 79% suburban settings. American teachers were 83% White, 13% African American, and 4% other ethnicities.

The 502 secondary school teachers were from Canada (n = 255) and Singapore (n = 247). Canadian secondary teachers were 54% female, with a mean age of 38.90 (SD = 11.07), 12.85 years of experience (SD = 10.22), and taught in public secondary schools (82%) or in other combinations of multi-graded settings that included secondary and junior high ages (18%). Most of the Canadian teachers were born in Canada (94%) and described ethnic heritage as Anglo-Canadian (47%), European Canadian (29%), French Canadian (7%), East Asian Canadian (2%) and other/no response (15%). Singaporean data collection was stratified to ensure that participants would represent teachers from the full range of school streams. Teachers from Singapore were 63% female with a mean age of 35.76 (SD = 9.70), 10.46 years of experience (SD = 10.29), and taught in secondary schools described as “government” (54%), “autonomous” (44%), and not indicated (2%). The sample included representatives from the highest ability “Special” stream (11%), high ability “Express” stream (28%), and the average and lower ability “Normal” and “Normal Technical” streams (33%), with 31% not identified. Most of the Singaporean teachers were born in Singapore (89%) or Malaysia (6%), and described ethnic heritage as Asian (98%).

We used convenience samples of teachers from a range of schools in each country. Participants in each setting were volunteers who were approached in their schools (in Cyprus, Singapore, and United States) or at conferences or workshops (Canada and Korea) and asked to fill out a brief questionnaire. Response rates

### 2. Method

#### 2.1. Participants and procedure

Participants were 1212 teachers from five countries teaching in elementary/middle schools (n = 709) and secondary schools (n = 502). We combined teachers from elementary and middle schools in each country, and the United States (participants from Cyprus were all elementary school teachers) into single groups because (a) elementary and middle school often serve overlapping student populations (e.g., K-7 and grades 6–8), (b) the two school structures typically include a stronger focus on building teacher-
2.2. Measures

The TSE measure contained 12 items, with four items in each of the three factors: TSE for instructional strategies, TSE for student engagement, and TSE for classroom management. Participants completed a 9-point response scale, ranging from 1, Nothing, to 9, A great deal to rank their self-efficacy related to 12 teaching-related tasks (e.g., “How much can you do to craft good questions for students?”). The 12 items are included in Table 2. Job satisfaction was measured with a single item taken from Caprara et al. (2003): In general, I am satisfied with my job, with responses options ranging from 1, Disagree strongly, to 9, Agree strongly. Singapore teachers did not complete the job satisfaction measure. Single-item measures of job satisfaction are frequently used in research exploring workplace settings where survey length and face validity are critical. Recent studies have found single-item measures to show high correlations with multiple-item measures (e.g., Dolbier, Webster, McCalister, Mallon, & Steinhardt, 2005; Nagy, 2002; Wanous, Reichers, & Hudy, 1997), and have concluded that single-item measures of job satisfaction are an acceptable and reasonable alternative to multiple-item scales.

In order to create Greek (for Cyprus) and Korean versions of the instruments that displayed linguistic, functional, and cultural equivalence, we followed translations protocols used in cross-cultural research. First, we used a team or committee approach to the translation-back translation process that allows for multiple checks on functional and cultural validity (Peña, 2007). Second, we used bi- or multi-lingual translators who were experts in the research domain thus ensuring that the translations were not only linguistically accurate, but valid in substance and meaning (Van de Vijver & Leung, 1997). Third, our translations were guided by a meaning-based approach in which changes in sentence structure and wording were allowed in the translated version to reflect differences in thought patterns and syntax differences between the original and translated version of the instrument (Larson, 1998). The Cyprus version of the questionnaire was translated into Greek by a fluently bilingual educational psychology researcher who is familiar with teacher motivation research, and back-translated by a Cypriot education specialist with a good knowledge of English and of educational psychology. Disagreements over wording were solved through discussion by the two Cypriot researchers, in consultation with the first author. The Korean version of the questionnaire was translated into Korean by the second author, who is fluent in English and Korean and who specializes in motivation research, and independently back-translated by three translators with excellent knowledge of Korean and English. The independent translations resulted in general linguistic agreement, with only minor differences in wording of the items. Missing data across all groups were minimal (less than 1%) and were replaced by group item means. Teachers in Canada, Singapore, and the United States completed the survey in English.

2.3. Plan of analysis

The main analyses of the study involved multi-group confirmatory factor analysis (MGCFA) using AMOS 7.0 (Arbuckle, 2006). The use of confirmatory factor analysis has been called the “method of choice” for cross-cultural and international comparisons through the assessment of measurement invariance across groups (Stein, Lee, & Jones, 2006, p. 249). We tested the invariance of the TSES with four increasingly stringent tests. First, because the TSES is a relatively new measure, we considered whether the measure was better conceptualized across all six groups as a one-factor or a three-factor construct. Second, we evaluated the TSES across the three North American groups (teachers from Canadian elementary/middle schools, American elementary/middle schools, and Canadian secondary schools) in order to investigate (a) if the measure operated in similar ways in Canada as in the United States and (b) if the measure operated in similar ways across elementary/middle school and secondary school settings within a North American context. Third, we compared the TSES across the two Asian groups (teachers from Korean elementary/middle schools and Singaporean secondary schools) in order to investigate (a) if the measure operated in similar ways across two Asian countries that may share some cultural beliefs and practices such as collectivism and Confucian values (Park & Huebner, 2005) and (b) if the measure operated in similar ways across elementary/middle school and secondary school settings within an Asian context. Finally, we examined the TSE measure across all six groups from five different countries to investigate if the measure displayed invariance across the full range of settings.

After testing the one-factor and initial three-factor models, three-factor baseline models were established separately in each group. In establishing baseline models, the basic factor structure was maintained for all groups, with only error covariances allowed to differ across groups. The baseline models were used to (a) conduct a simultaneous test of equal factor structure across all groups (configural invariance), (b) test the equality of factor loadings across groups (construct-level metric invariance), (c) test the equality of structural relations among the factor variances (invariance in the range of responses given to each item), and (d) test the equality of structural relations among factor covariances (invariance in the relationship among latent constructs). Hierarchical goodness-of-fit was evaluated by the conventional use of change in $\chi^2 (\Delta \chi^2$), in which a non-significant $\Delta \chi^2$ when constraints are imposed indicates invariance, and by changes in the comparative fit index (ACFI), wherein an ACFI less than or equal to .01 indicates invariance when additional constraints are imposed (Cheung & Rensvold, 2002). The use of $\Delta$CFI is superior to the use of $\Delta \chi^2$ as a test of invariance because $\Delta$CFI is not affected by sample size and is preferable to the use of other goodness-of-fit change indices that possess undesirable qualities that influence the accuracy of measurement when invariance constraints are added (see Cheung & Rensvold, 2002, for further explanation).

Internal consistency coefficients (alpha) were calculated, and levels of the TSE composite and subscales were compared across groups using ANOVA, with Scheffé post hoc analyses used to identify mean differences between groups. To test for further evidence of construct validity, we explored the relationship between the TSES and a single-item job satisfaction measure across the five groups (excepting Singapore where job satisfaction data were not collected). We hypothesized that TSE would be positively associated with job satisfaction in each setting as has been reported previously in single-culture studies (e.g., Caprara et al., 2003, 2006). Similar patterns of correlations between job satisfaction and TSE...
3. Results

Across the five groups would provide additional support for the cross-national construct validity of the TSEs. Because social cognitive theory is hypothesized to operate in similar ways across cultures (Bandura, 2002), we predicted similar correlations between job satisfaction and the TSE across groups.

3.1 One-factor TSE models

The factor loadings were similar across all groups, with loadings ranging from .75 to .87. The factor variances were also similar, with values ranging from .64 to .74. The factor covariances tended to be low, with values ranging from -.02 to .02. The fit indices for the one-factor model were generally good, with CFI values ranging from .97 to .99 and RMSEA values ranging from .04 to .07.

3.2 Three-factor initial TSE models

The fit indices for the three-factor model were generally good, with CFI values ranging from .95 to .98 and RMSEA values ranging from .04 to .06. The factor loadings were similar across all groups, with loadings ranging from .55 to .80. The factor variances were also similar, with values ranging from .61 to .71. The factor covariances tended to be low, with values ranging from -.02 to .03.

3.3 Three-factor baseline TSE models

The fit indices for the three-factor baseline model were generally good, with CFI values ranging from .95 to .98 and RMSEA values ranging from .04 to .06. The factor loadings were similar across all groups, with loadings ranging from .55 to .80. The factor variances were also similar, with values ranging from .61 to .71. The factor covariances tended to be low, with values ranging from -.02 to .03.

Table 1 presents the fit indices for the one-factor, initial three-factor, and baseline three-factor models for the six individual groups, and across combinations of groups (North American, East Asian, and all six groups) for TSE. We included three measures of goodness-of-fit: $\chi^2/df$ ratio, wherein a ratio $<3.0$ indicates a good fit; CFI, where $>0.90$ indicates a good fit, and root mean square error of approximation (RMSEA), where $.05$ is the upper limit of acceptability. Measurement invariance was assessed by imposing cross-group constraints on the measurement model.

A one-factor model of TSE was compared to the initial three-factor model proposed by Tschannen-Moran and Woolfolk Hoy (2001). The one-factor model showed a poor fit for all groups, with $\chi^2/df$ ratios ranging from 4.92 for Korean teachers to 8.03 for Canadian secondary teachers, and CFIs ranging from .640 in the Canadian secondary group to .870 in the Singaporean group. A three-factor model initial model was a significant improvement over the one-factor model in each setting, and for the combined groups, $\Delta \chi^2(18) = 1176.91, p < .001$. The $\chi^2/df$ ratios for the three-factor initial model ranged from 1.38 in the American group to 3.71 in the Korean group, with CFIs ranging from .974 in the American group to .876 in the Korean group.

In establishing the baseline models, the basic factor structure was maintained, but following the direction of Byrne and colleagues (e.g., Byrne, 2004; Byrne, Shavelson, & Muthén, 1989), we added group-specific correlated error terms based on the modification indices, item-scale characteristics (e.g., adjacent items on a scale), and on similarities in wording among the items. The same correlated error terms ($\Delta_{11,12}$) were added to the two Canadian groups, the Korean group, and the Singaporean group and resulted in significantly improved fit for the Korean teachers, an additional factor model.
correlated error term (δ_{1,3}) was added, and for the Cyprus group, two unique correlated error terms were added (δ_{6,9}, δ_{5,10}). No correlated error terms were deemed necessary in the American group.

The baseline models for the individual groups showed good fit, with all χ^2/df ratios below the threshold proposed by Byrne (2004), ranging from 1.38 for the American group to 2.66 in the Korean group. The CFIs ranged from .927 in the Korean group to .977 in the Singapore group, all above the minimum CFI threshold proposed by Byrne (2004). The RMSEAs ranged from .051 in the Canadian elementary group to .104 in the Korean group, with all RMSEA levels except for the Korean group below the .09 upper limit described by Byrne (2004). Before testing for invariance across all groups, invariance was assessed across samples of elementary/middle school and secondary teachers from North America (i.e., from Canada and the United States), and subsequently in samples of elementary/middle school and secondary teachers in East Asia (i.e., Korea and Singapore). Testing for invariance within culturally similar regions serves two purposes in our study—first, it provides a comparison for the results from the mixed-culture, and second, it provides a test for validity of the TSES when comparing responses of teachers from different school levels. The North American and East Asian groups included teachers from elementary/middle schools and secondary schools; invariance across each of these similar-culture groups argues for the validity of the TSES in both settings.

3.1. Invariance across North American groups of teachers

The unconstrained model (consisting of the baseline models for the three North American groups) showed a very good fit, χ^2/df = 1.53, CFI = .973, RMSEA = .030, suggesting a common factor structure across the three groups. Constraining the pattern coefficients across the three groups resulted in continued excellent fit, Δχ^2 of 19.90 (Δdf = 18), p = ns, ΔCFI = .001, suggesting that factor weights were invariant across the three groups. Constraining structural variances yielded non-significant changes in fit, Δχ^2 of 1.83 (Δdf = 6), p = ns, ΔCFI = .001, suggesting no difference in structural variances across the three groups. Finally, constraining structural covariances resulted in no changes in goodness-of-fit, Δχ^2 of 6.04 (Δdf = 6), p = ns, ΔCFI = .000. Invariance testing across the three North American groups of teachers suggested that the TSES operates in a similar way for these groups of teachers.

3.2. Invariance across East Asian groups of teachers

The invariance of the TSES was measured for the groups of elementary/middle school teachers from Korea and secondary teachers from Singapore. The unconstrained model showed a very good fit across groups, χ^2/df = 2.33, CFI = .960, RMSEA = .058. Constraining the pattern coefficients across the two groups resulted in a non-significant change in chi-square, with little change in the CFI, Δχ^2 of 15.30 (Δdf = 9), p = ns, ΔCFI = .002. Constraining structural variances across groups yielded a non-significant change in chi-square, with little change to the CFI, Δχ^2 of 2.69 (Δdf = 3), p = ns, ΔCFI = .002. Constraining structural covariances across groups yielded non-significant changes, Δχ^2 = 1.48 (Δdf = 3), p = ns, ΔCFI = .000. These results suggest invariance of form, factor loadings, and factor variances and covariances of the TSES in two Asian settings.

3.3. Invariance across six groups of teachers

The multigroup analysis with no constraints resulted in a baseline χ^2 value of 588.47, df = 299, p < .001, with very good fit (CFI = .979, χ^2/df = 1.97, and RMSEA = .028), providing support for a common three-factor structure across the six groups. Constraining the factor loadings across all groups resulted in a significant Δχ^2 of 114.42 (Δdf = 45), p < .001, with an acceptable ΔCFI (.009), falling within the .01 limit proposed by Cheung and Rensvold (2002). Constraining structural variances across groups yielded a significant change in chi-square, Δχ^2 of 35.76 (Δdf = 15), p = .01, but with negligible ΔCFI of .002. Constraining structural covariances across groups yielded a significant change in chi-square, Δχ^2 of 51.08 (Δdf = 15), p < .001, but again with minimal ΔCFI of .004. These results suggest that the TSES shows measurement invariance across the six groups, with invariance of form, factor loadings, with mixed support for invariance of factor variances and covariances.

Table 2 presents standardized factor pattern coefficients (jis) and interfactor correlations (jos) for the 12 individual items of the TSE scale and for the three factors of self-efficacy for instructional strategies, self-efficacy for student engagement, and self-efficacy for classroom management. All pattern coefficients displayed adequate to high factor loadings. Item 11—How much can you do to assist families in helping their children do well in school?—displayed the lowest factor loading in five of six settings. The three TSE factors were significantly correlated in each setting.

3.4. Levels of teachers’ self-efficacy and job satisfaction

Table 3 presents reliability indexes, means, and standard deviations for the TSE composite scale, three subscales, and means and standard deviations for the job satisfaction measure. Cronbach’s alpha reliability indices ranged from .71 for the TSE for instructional strategies subscale in the Canadian secondary teacher sample, to .94 for the TSE composite score in the Singaporean sample. The overall pattern of results showed participants from the two Asian settings—Korea and Singapore—to rate TSE lower than participants from the four other groups. Levels of the composite TSE scores varied across groups, F(5,1211) = 55.45, p < .001, η^2 = .19. Post hoc comparisons using Scheffé multiple comparisons showed no differences between the Korean and Singaporean teachers (p = .41), but significant differences between the two Asian groups and each of the other four groups (all ps < .001). Levels of TSE for instructional strategies varied across groups, F(5,1211) = 70.32, p < .001, η^2 = .23. Post hoc comparisons showed Korean teachers with significantly lower scores than Singaporean teachers, who in turn scored lower than teachers in the other four settings. Levels of TSE for student engagement varied across groups, F(5,1211) = 25.35, p < .001, η^2 = .10. Post hoc comparisons showed ratings from Korean, Singaporean, and Canadian secondary teachers to be similar, and significantly lower than Canadian elementary, Cypriot, and American teachers. Levels of TSE for classroom management varied across groups, F(5,1211) = 51.67, η^2 = .18. Post hoc comparisons showed no differences between the two Asian groups, which in turn were both significantly lower than the four other groups. Levels of job satisfaction did not vary across groups, F(4,965) = 1.47, p = .27.

3.5. Correlations

Table 4 presents the correlations between job satisfaction and TSE for five groups (job satisfaction data were not collected in Singapore). The TSE composite score and each of the subscales were significantly related to job satisfaction in each setting. The direction and size of correlations were similar between settings, with all correlations positive, ranging from r = .17 (p < .05) for the correlation between job satisfaction and TSE for instructional strategies in the Korean sample, to r = .48 (p < .001) for the correlation between job satisfaction and TSE for instructional strategies in the Cyprus sample. Comparisons using Fisher’s z-transformations revealed similar magnitude of correlations in 37 of 40 possible comparisons. One
significant difference was found between Korean and Cypriot teachers on the relationship between job satisfaction and TSE for instructional strategies \((p = .003)\), and two differences were found between job satisfaction and TSE for classroom management between American teachers and Canadian secondary teachers, \((p = .02)\) and between American and Cypriot teachers \((p = .01)\). There were no between-group differences in the correlations between job satisfaction and the TSE composite, or between job satisfaction and TSE for student engagement.

4. Discussion

Research on teachers’ sense of self-efficacy has been growing over the last decade, but little attention has been paid to building understanding of the universality of the construct in diverse settings. Although previous TSE research has been hampered by flawed conceptualization, the measure proposed by Tschannen-Moran and Woolfolk Hoy (2001) addresses many of these conceptual flaws. This article explored the cross-national validity of the Teachers’ Sense of Efficacy Scale (TSES) across six groups in five countries and extends our understanding of this “little idea with big impact” (Tschannen-Moran & Woolfolk Hoy, 2007, p. 954). We found that the TSES showed convincing evidence of invariance of factor forms, factor loadings, and factor variances and covariances across groups of teachers within culturally similar regions in North America and East Asia, and across six groups of teachers from five countries in North America, East Asia, and Europe. Although little research has explored how TSE operates in diverse settings, the results from this study show that items on the TSES show strong internal consistency not only in American settings, but also in Canadian, Cypriot, Korean, and Singaporean settings. Similar correlations between TSE and job satisfaction were also found across the different groups, as hypothesized, providing further evidence of international validity.

The TSE measure showed strong measurement invariance in groups of teachers from similar-culture geographical regions (i.e., Canadians and Americans in North America; Koreans and Singaporeans from East Asia) even though the teachers came from different levels of schools. Strong measurement invariance was found for groups of teachers from Canada and the United States who shared similar languages, cultural dimensions, and school systems, but differed on school level, teacher preparation, and teaching environment. Similarly the TSES showed strong measurement invariance within the two Asian groups. The groups of teachers from Korea and Singapore share some cultural practices and beliefs (e.g., similar levels of individualism and collectivism; Hofstede & Hofstede, 2005), and may share some educational values and practices, but the two groups differed in school level, and in language, geography, history, and some cultural practices. The TSES also showed evidence of invariance when all six groups from five countries were compared simultaneously. In spite of the significant differences in the hierarchical chi-square tests, there was support for measurement and structural invariance from the tests using ACFI, which has been recommended as preferable to the use of \(\Delta \chi^2\) as a test of measurement invariance when conducting multigroup CFA (Cheung & Rensvold, 2002). Researchers who are interested in the comparative study of teachers’ motivation beliefs across cultures and countries should consider the TSES as a measure that displays a strong conceptual and theoretical foundation, coupled with acceptable cross-settings invariance.

Although there were minor variations, the levels of TSE showed a pattern of higher scores in Canada, Cyprus, and the United States, with relatively lower scores in Korean and Singaporean samples of teachers. Similar patterns of results held true for the TSES composite score, as well as for the TSE for instructional strategies and TSE for classroom management across groups, with the two Asian
Table 3

<table>
<thead>
<tr>
<th>Country</th>
<th>TSE for instructional strategies</th>
<th>TSE for student management</th>
<th>TSE for classroom management</th>
<th>TSE total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Canada (elem/middle) n = 210</td>
<td>0.78 (0.06)</td>
<td>0.85 (0.06)</td>
<td>0.80 (0.06)</td>
<td>0.85 (0.06)</td>
</tr>
<tr>
<td>Cyprus n = 210</td>
<td>0.77 (0.06)</td>
<td>0.85 (0.06)</td>
<td>0.85 (0.06)</td>
<td>0.86 (0.06)</td>
</tr>
<tr>
<td>Korea n = 210</td>
<td>0.88 (0.06)</td>
<td>0.84 (0.06)</td>
<td>0.85 (0.06)</td>
<td>0.89 (0.06)</td>
</tr>
<tr>
<td>Singapore (secondary) n = 247</td>
<td>0.85 (0.06)</td>
<td>0.85 (0.06)</td>
<td>0.85 (0.06)</td>
<td>0.87 (0.06)</td>
</tr>
<tr>
<td>United States (elem/middle) n = 137</td>
<td>0.89 (0.06)</td>
<td>0.80 (0.06)</td>
<td>0.85 (0.06)</td>
<td>0.89 (0.06)</td>
</tr>
<tr>
<td>United States (secondary) n = 201</td>
<td>0.79 (0.06)</td>
<td>0.84 (0.06)</td>
<td>0.84 (0.06)</td>
<td>0.87 (0.06)</td>
</tr>
</tbody>
</table>

Note: Job satisfaction data were not collected in Singapore.
man functioning in a wider range of cultural settings (Pajares, 2007). There are five important results emerging from this study. First, results confirmed that a three-factor model of TSE consisting of self-efficacy for student engagement, instructional strategies, and classroom management, is statistically preferable over a one-factor model of TSE in a range of settings. Second, results from this study show that items on the TSES demonstrate internal consistency in a variety of settings. Third, findings from this study show that the TSES shows measurement invariance across groups of teachers in similar cultural groups but from different school types, namely, elementary/middle schools and secondary schools. Fourth, the study presents evidence of measurement invariance of the TSES across diverse cultural groups, including groups of teachers from North America, Asia, and Europe. Finally, the study provides support for similar relationships of TSE with job satisfaction across six groups in five countries.

The findings from this study have implications for educational practice and pave the way for future research using the TSES in diverse cultural contexts. Previous studies have shown that TSE influences student achievement and motivation as well as teachers’ classroom behaviors and motivation. The current study confirms that each of the three factors of the TSES is positively related to teachers’ job satisfaction in each of the settings explored, with consistent correlations between the TSES total and job satisfaction in elementary/middle school and secondary samples in five countries. The results from the study suggest that teachers with high levels of self-efficacy experience greater job satisfaction in a wide variety of school and national settings. Although the direction of the TSE-job satisfaction relationship was not explored in this study, it is plausible that strengthening teachers’ self-efficacy beliefs through targeting the sources of self-efficacy will enhance the satisfaction teachers gain from their work. School administrators would do well to focus on building teachers’ self-efficacy by providing clear opportunities for successful experience, positive modeling from successful peers, and verbal encouragement. For novice teachers who have not had adequate opportunity to build successful experience, and for whom self-efficacy may be most malleable, positive modeling and verbal encouragement may be especially important in building self-efficacy (Tschannen-Moran & Woolfolk Hoy, 2007).

Job satisfaction is only one of the hypothesized outcomes of the self-efficacy of teachers; student motivation and performance may also be enhanced when teachers become more confident about their capabilities to successfully engage students in a variety of challenging circumstances. Future cross-national studies should explore the relationship of TSE with student outcomes, like academic performance and motivation. Perhaps the most exciting implication of this study for educational practice is that the TSES showed itself to be a valid tool to measure the self-efficacy beliefs of teachers in a range of settings. Studies can now be conducted to explore a wide variety of educational outcomes in a wider range of settings than was previously possible. The current study adds to our understanding of the motivation beliefs of teachers, and provides support for the use of the TSES outside of culturally Western settings. Although teachers’ personal beliefs and teaching conditions may contrast according to cultural background and education system, beliefs about personal capabilities to teach appear to operate in similar ways across a diverse range of countries. Moreover, the relationship between teacher efficacy and beliefs about the satisfaction derived from teaching appears to function in the same way across borders.

Acknowledgments

The authors gratefully acknowledge funding support to the first author from the Social Sciences and Humanities Research Council of Canada. Thanks are also extended to the anonymous reviewers for suggestions that were most helpful in improving the clarity of the paper.

References


