Article:

Burić, Irena and Kim, Lisa orcid.org/0000-0001-9724-2396 (2021) Job satisfaction predicts teacher self-efficacy and the association is invariant: Examinations using TALIS 2018 data and longitudinal Croatian data. Teaching and Teacher Education. 103406. ISSN 0742-051X

https://doi.org/10.1016/j.tate.2021.103406

Burić, I. & Kim, L. E. (in press) Job satisfaction predicts teacher self-efficacy and the association is invariant: Examinations using TALIS 2018 data and longitudinal Croatian data. *Teaching and Teacher Education*. Advance online publication. doi: 10.1016/j.tate.2021.103406

Job satisfaction predicts teacher self-efficacy and the association is invariant:

Examinations using TALIS 2018 data and longitudinal Croatian data

Irena Burić[a] & Lisa E. Kim[b]

[a] Department of Psychology, University of Zadar, Croatia

[b] Department of Education, University of York, UK

Corresponding author:

Irena Burić
Department of Psychology, University of Zadar,
Obala kralja Petra Krešimira IV., br. 2
23000 Zadar
Croatia

E-mail: iburic@unizd.hr

Funding: This research was supported by Croatian Science Foundation (Grant No. UIP-11-2013-5065)

Highlights

- The association between teacher self-efficacy and job satisfaction is invariant across teachers' gender, career stage, instructing educational level, and time
- Job satisfaction predicts future teacher self-efficacy
- Teacher self-efficacy does not predict future job satisfaction

Abstract

Theoretical models and empirical studies have long assumed that the positive association between teacher self-efficacy (TSE) and job satisfaction (JS) is invariant regardless of any teacher factors and that TSE predicts JS. Using OECD TALIS 2018 data and a three-time point longitudinal Croatian data, we test these assumptions by examining the invariance of the association between TSE and JS across four factors (gender, career stage, educational levels, and time) and the causal ordering of the two constructs. We find support for the first assumption but not the second; that is, JS predicts TSE and TSE does not predict JS.

Keywords: teacher self-efficacy; job satisfaction; TALIS; structural equation modelling

Job satisfaction predicts teacher self-efficacy and the association is invariant: Examinations using TALIS 2018 data and longitudinal Croatian data

Teacher shortage is an international problem (Education for All Global Monitoring Report and the UNESCO Education Sector, 2015; IBF International Consulting, 2013). In fact, the world needs almost 69 million new teachers by 2030 to provide every child with primary and secondary school education (UNESCO Institute for Statistics, 2016). Unfortunately, many countries are experiencing a teacher shortage due to high attrition rates (Dupriez et al., 2016), which are as high as 50% for teachers in the first five years of the profession (Buchanan et al., 2013; Ingersoll, 2003; Lindqvist et al., 2014). Some countries have attempted to address this problem by introducing new policies and support services to encourage teachers to remain in the profession (e.g., Department for Education, 2019; NSW Government, 2020). Strategies aimed to encourage retention have often been aimed to increase teachers' job satisfaction (JS) and/or self-efficacy (TSE), especially given that these are one the most studied teacher wellbeing and motivation constructs with well-established associations with a variety of important measures, such as turnover intentions and job commitment (see Gkolia et al., 2014; Zee & Koomen, 2016 for reviews).

To ensure maximum applicability and effectiveness in allocations of attention and resources to decrease attrition rates and improve educational quality, one must first ascertain the nature of the association between the relevant constructs. Two widely-held assumptions pervade the TSE and JS literature in regard to the nature of their association: (a) although the mean levels of TSE and JS seems to differ across teacher factors, such as gender and instructing educational level (Klassen & Chiu, 2010; Viel-Ruma et al., 2010), the strength of the TSE–JS association is similar (see Zee & Koomen, 2016 for a review); and (b) TSE precedes JS in its causal order (see Gkolia et al., 2014 for a review). However, these assumptions have not yet been empirically interrogated. Given the critical roles of TSE and

JS in the lives of both teachers and students (OECD, 2020), it is imperative that the absence of such a study is addressed (Schleicher, 2018). Accordingly, we use two large independent datasets (i.e., OECD Teaching and Learning International Survey 2018 data and three-wave longitudinal teacher dataset) to examine, what we believe is for the first time, the invariance of the association between TSE and JS across gender, career stage, educational level, and time, as well as the causal ordering of the two constructs.

Invariance of the TSE–JS Association

TSE is one of the most studied motivation constructs, which can be defined as one's belief in their ability to carry out various aspects of their job (Tschannen-Moran et al., 1998). More specifically, TSE refers to teachers' judgements that they are able to implement desired instructional strategies, manage classroom effectively, and engage students in learning (Tschannen-Moran & Hoy, 2001). This construct is grounded in Bandura's (1986, 1997) social-cognitive theory, which posits that self-efficacy is influenced by behavioral, environmental, affective, and cognitive factors and is necessary to fuel one's persistent effort to reach their goals. TSE has often been viewed as a predictor of a myriad of job-related constructs associated with classroom processes, student academic adjustment, and teacher wellbeing (e.g., Author, 2018; 2020a; Collie et al., 2012).

JS refers to one's experiences of pleasurable emotional states derived from their evaluation of their job (Locke, 1969). However, JS can also be understood as a multidimensional psychological response to the cognitive (i.e., evaluative) and affective (i.e., emotional) component of their job (Judge, Hulin, & Dalal, 2009). It can be assessed either as (a) an overall or a global evaluation of one's job (Weiss, 2002) or (b) as a multifaceted construct (e.g., satisfaction with work, satisfaction with pay, satisfaction with conditions; Spector, 1997). Organizational psychologists have particularly taken interest in this construct given its established meta-analytic associations with a variety of important outcomes, including organizational citizenship behavior (Organ & Ryan, 1995) and job performance (Judge et al., 2001). In the teaching profession, JS has been found to be negatively associated with intention to quit (Carson et al., 2017) and positively with teacher–student relationships (Veldman et al., 2013).

A systematic review found that the association between TSE and JS was relatively stable across 21 studies, with a median coefficient of .33 (Zee & Koomen, 2016). That is, primary and secondary school teachers with higher levels of self-efficacy were also more satisfied with their jobs and their relationships in their jobs. Apparent in the study findings, however, is the assumption that the nature of the TSE–JS association is the same regardless of individual teacher factors. Such an assumption has not yet been empirically examined. To ensure that research and practice most accurately represent and implement the true nature of their association, one must examine the robustness and the generalizability of the TSE–JS association across common moderators, such as gender, career stage, teachers' instructing educational level, and time.

Findings from empirical studies suggest that the TSE–JS association may indeed be invariant across these factors. For example, although male teachers have slightly higher mean levels of TSE and JS than females (e.g., Crossman & Harris, 2006; Skaalvik & Skaalvik, 2007), gender is not correlated with JS (Klassen & Chiu, 2010). In regard to career stages, a study found that the associations between indicators of teachers' professional identity, which included TSE and satisfaction with salary and relationships, were invariant across novice, experienced, and senior teachers (Canrinus et al., 2012). Regarding the factors of educational levels and time, mean levels of TSE and JS seems to differ across educational levels (Klassen & Chiu, 2010; Viel-Ruma et al., 2010) but there is no evidence on time nor the invariance of the TSE–JS association across these two factors. However, as a study reported the invariance across these factors in regard to the association TSE has with burnout, which a contextualized motivation variable like JS (Author, 2020b), findings of which suggests that TSE–JS association may also be invariant across educational level and time. Thus, although studies have not explicitly examined the invariance of the TSE–JS association, the association may in fact be invariant across gender, career stage, teachers' instructing educational level, and time.

Causal Ordering of TSE and JS

Another common assumption pervading TSE and JS literature is their causal ordering. Although the constructs' ordering has been represented in a myriad of ways, empirical studies have predominantly depicted TSE as a predictor of JS (e.g., Caprara et al., 2006; Skaalvik & Skaalvik, 2017; Wang et al., 2015). However, a minority of studies have represented the two constructs as being at the same level; that is, they were either both predictors (e.g., Adebomi & Olufunke, 2012; Soto & Rojas, 2019), outcome variables (e.g., Mottet et al., 2004; Taylor & Tashakkori, 1995), or reciprocal correlates (e.g., Avanzi et al., 2013). To authors' knowledge, no studies have examined whether JS predicts TSE and/or whether TSE and JS have cross-lagged bidirectional associations.

An added complication to the accepted belief on the constructs' causal ordering is that it has been inferred from correlational data. In fact, Zee and Koomen (2016) reported in their systematic review that although 19 of the 21 studies reporting the TSE–JS association were cross-sectional, most had inferred causality by using statistical analyses beyond correlations, such as structural equation modelling. Nevertheless, Zee and Koomen proposed that it is likely that TSE may be associated with teacher attrition indirectly through factors such as JS. Such a proposal is not misaligned with most theoretical models, which outlines that motivation variables are determinants of measures of occupational wellbeing and effective teaching practices (e.g., Kunter et al., 2013; Richardson et al., 2014). For example, according to the model of job stress and satisfaction in teachers (Troesch & Bauer, 2017), TSE influences one's positive and negative appraisals of the job, which then predict JS. That is, self-efficacy may influence one's job-related thoughts and behaviors that contribute to one's performance and interactions with people, which then predicts JS (Akomolafe & Ogunmakin, 2014).

Other models, however, propose alternative ways that TSE and JS may be ordered. According to the social-cognitive theory (Bandura, 1997), satisfied individuals tend to believe they are more self-efficacious. For example, according to Bandura (1997), self-efficacy beliefs are formed based on four groups of factors: enactive mastery experiences, vicarious experiences, social persuasion, and physiological and affective states. Considering that affect is in the core of JS (Judge et al., 2001; Locke, 1996), teachers who are more satisfied with their jobs also experience more positive affect, which consequently leads to more positive self-efficacy beliefs. In turn, self-efficacy beliefs shape individuals' affective reactions to the task (Gist & Mitchell, 1992), implying that TSE can influence JS as well. Next, the broaden-and-build theory (Fredrickson, 2004) offers another view on the possible direction of the association between TSE and JS. According to the theory, positive emotions broaden individuals' thought–action repertoire, which in turn can help build personal resources such as TSE. Since JS encompasses an emotional component, satisfied teachers may possess broader scopes of attention and cognition as well as to think in more flexible and creative way, which may make them more efficacious when faced with different demands of their job.

Lastly, OECD's Model of Teachers' Wellbeing and Quality Teaching (Schleicher, 2018) situates TSE and JS in the wellbeing dimensions of the model but does not specify the nature of the association between them, perhaps alluding to how they may be closely interlinked with each other. Additionally, Zee and Koomen (2016) noted in their heuristic model that there may be a reciprocal association between TSE and wellbeing factors, including JS, but that it is impossible to ascertain this as there are no studies examining this.

The absence of studies examining the possible reciprocal association and the antecedent order of the two constructs may be due to the strict requirements the data must fulfil in order to conduct such analyses. For example, a fundamental requirement is that the data must be longitudinal (Hamaker, 2005), which only two of the studies on TSE and JS were (see Zee & Koomen, 2016 for a review). However, one of these studies measured TSE at two time points (6 months apart) but JS at only one time point (Avanzi et al., 2013). Another study measured both TSE and JS at two time points but their statistical model only assessed the association between TSE and JS within each of the two time points (Salanova et al., 2011). That is, they did not test whether either of the constructs at T1 predicted the other construct at T2. Thus, there are yet no empirical studies, from the authors' knowledge, which have examined different permutations of the possible associations TSE and JS may have with each other. Accordingly, the current study also examines whether TSE predicts JS, JS predicts TSE, and/or whether they are reciprocal correlates.

Overview of the Current Studies

This paper consists of two studies aimed to thoroughly explore the nature and robustness of the association between TSE and JS. In Study 1, we use OECD Teaching and Learning International Survey (TALIS) 2018 data (OECD, 2019b) to establish the cross-sectional association between TSE and JS on a large sample of teachers from diverse countries (N = 252,881) and to test its association across important possible moderators; that is, across gender, career stage, and educational level. In Study 2, we use data obtained from a longitudinal study conducted on a large sample of Croatian teachers (N = 3,002) to uncover the temporal order between the two constructs and to test the invariance of established longitudinal relations across the moderators; that is, gender, career stages, educational level, and time.

It should be noted that even though both studies examined the same constructs of TSE and JS, their operationalization slightly differed, which can be helpful in assessing the generalization of the findings regardless of their specific conceptualization. More specifically, in Study 1, TSE and JS were measured as multidimensional constructs (OECD, 2019b). TSE comprised of three dimensions (i.e., self-efficacy in classroom management, self-efficacy in instruction, and self-efficacy in student engagement), which is in line with Tschannen-Moran and Hoy's (2001) definition of TSE. JS was measured by two dimensions: JS with work environment (i.e., school in which a teacher works) and JS with profession in more general. In Study 2, TSE was assessed as a unidimensional construct comprised of teachers' judgments about their job accomplishment, skill development, social interaction with students, parents, and colleagues, as well as coping with job stress, consistent with other literatures' conceptualization of TSE (e.g., Schwarzer et al., 1999). In addition, JS was measured as an overall affective evaluation of one's job (Judge et al., 1998). Despite these differences, both measures of TSE are strongly rooted in Bandura's (1986, 1997) social-cognitive theory and assess teachers' beliefs in their ability to successfully deal with various aspects of their job. Similarly, both measures of JS grasp teachers' evaluative and emotional responses to their job that are in the core of job satisfaction definition (Judge et al., 2009). Therefore, combining the two complementary data sets based on diverse samples (i.e., international vs. national), methodologies (i.e. cross-sectional vs. longitudinal), and somewhat different operationalization of constructs (i.e., multi-faceted vs. global), allows us to gain a more comprehensive insight into the true nature of the associations and their generalizability.

Study 1

Although studies have found that TSE is positively associated with JS for teachers in countries such as Finland (Malinen & Savolainen, 2016) and Canada (Collie et al., 2012), studies have not yet examined this link for teachers using multiple countries. TALIS is the largest international survey on teachers, assessing a wide range of aspects of their professional lives, including TSE and JS (OECD, 2019a). As such, we use this unique dataset to examine whether the TSE–JS association varies across (a) gender, (b) career stages (i.e., early-, mid-, vs. late-career), and (c) educational level the teacher instructs in. For reasons outlined above, we hypothesize that the TSE–JS association will be positive and invariant across these three factors (H1).

Methods

Participants and Procedure

The OECD has run TALIS once every five years since 2008 (Ainley & Carstens, 2018), with the most recent wave (TALIS 2018) taking place between September 2017 and July 2018. Teachers and school leaders from 48 OECD countries/economies responded to questions on various aspects of their professional lives (OECD, 2019a). We used country cases that OECD adjudicated as valid (OECD, 2019b) and contained data on both TSE and JS. As such, the study uses data from 46 countries/economies, which consisted of 252,881 unique teacher responses (68% females).

The teachers reported a wide range of age groups: under 25 (1.93%), 25-29 (9.45%), 30-39 (29.29%), 40-40 (31.04%), 50-59 (21.79%), 60 and above (6.26%), and some were missing (0.23%). Teachers ranged in their years of teaching experience (M = 16.17, SD = 10.33), which were then reclassified into career stages following the categories from Gu and Day (2007). That is, early-career (≤8 years of teaching experience; 27.39%), mid-career teachers (9-23 years of teaching experience; 47.21%), late-career teachers (≥24 years of teaching experience; 24.45%), and some were missing (0.95%). Moreover, teachers instructed in various International Standard Classification of Education (ISCED) levels:

Level 1 (18.68%), Level 2 (59.19%), and Level 3 (15.06%)[1]. According to the UNESCO Institute for Statistics (2011), ISCED 1 refers to level of primary education, ISCED 2 refers to level of lower secondary education, and ISCED 3 refers to level of upper secondary education.

Measures

The TALIS 2018 teacher questionnaire measures TSE using twelve items, all of which share the stem "In your teaching, to what extent can you do the following?" and the four-point Likert-scale, ranging from 1 *(not at all)* to 4 *(a lot)*. Three domains of self-efficacy are each measured by four items: self-efficacy in classroom management (e.g., "Get students to follow classroom rules"), self-efficacy in instruction (e.g., "Use a variety of assessment strategies"), and self-efficacy in student engagement (e.g., "Help students value learning"). The stratified Cronbach's alphas are .87 or higher across all countries and ISCED levels (OECD, 2019b). Items of all scales are shown in Appendix.

The TALIS 2018 teacher questionnaire measures JS using eight items, all of which share the stem "We would like to know how you generally feel about your job. How strongly do you agree or disagree with the following statements?". Participants responded on a four-point Likert scale, ranging from 1 *(strongly disagree)* to 4 *(strongly agree)*. An example item is "I enjoy working at this school." Two domains of JS are each measured by four items: job satisfaction with work environment (e.g., "All in all, I am satisfied with my job"), and job satisfaction with the profession (e.g., If I could decide again, I would still choose to work as a

[1] The ISCED levels for the TALIS PISA link data (7.07%) were not available. This could not be deduced from PISA data as teachers instructing 15 year olds in the TALIS PISA countries instruct in ISCED level 2, 3 or both (OECD, 2019c). Thus, TALIS PISA link cases were excluded when analyzing educational levels.

teacher"). The stratified Cronbach's alphas are .74 or higher across all countries and ISCED levels (OECD, 2019b).

Data Analyses

Statistical analyses were conducted in several steps. *First*, the Pearson correlation coefficients were calculated based on sum of scores on items measuring different domains of TSE and JS. In addition, due to the hierarchical structure of the TALIS data (i.e., teachers were nested in schools and schools were nested in countries), intraclass correlation coefficients (ICC1) were calculated to determine the amount of scales' variability that occurs at higher levels of analyses. ICC1 values greater than .05 (LeBreton & Senter, 2008) suggest that the analyses should be accommodated to account for the hierarchical nature of the data. *Second*, in acknowledgement of the multi-faceted nature of the measures of TSE and JS in Study 1, we conducted confirmatory factor analysis (CFA) to establish the baseline five-factor measurement model consisted of three TSE domains (i.e., self-efficacy in classroom management, self-efficacy in instruction, and self-efficacy in student engagement) and two JS domains (i.e., job satisfaction with work environment and job satisfaction with profession). Items measuring each of the domains served as indicators of their respective factors. *Third*, to establish the invariance of the association between TSE and JS across three moderators (i.e., gender, career stage, and educational level), a series of multi-group confirmatory factor analyses (MG-CFA) were conducted. More specifically, after establishing measurement invariance (i.e., configural and metric invariance; Millsap, 2001), covariance invariance across three moderators (i.e., gender, career stage, and educational level) was tested.

All analyses (except descriptive and correlational) were conducted using *Mplus 8.0* (Muthén & Muthén, 2017) using the robust maximum likelihood estimator (MLR). Missing data were handled by the full information maximum likelihood estimator (FIML; Enders, 2010), which is implemented in *Mplus* by default. Overall quality of model fit was evaluated based on several criteria: comparative fit index (CFI), Tucker-Lewis index (TLI), root-mean-square error of approximation (RMSEA), and standardized root-mean residual (SRMR). CFI and TLI values higher than .90 indicate an acceptable fit (Hu & Bentler, 1999), while RMSEA values lower than .06 and SRMR values lower than .08 indicate a good fit (Browne & Cudeck, 1993).

To compare the fit of the nested models, we used the criteria of ΔCFI < .010 and ΔRMSEA < .015, whereby lower values were preferred (Chen, 2007; Cheung & Rensvold, 2002). We did not use chi-square difference tests ($\Delta\chi 2$) as they are often statistically significant for models with large sample sizes, rendering the index uninformative (Dimitrov, 2010; Rudnev et al., 2018).

Results and Discussion

Descriptive Statistics, Pearson Correlations, and ICCs

We examined the descriptive statistics for analyzed variables and Pearson correlations between the demographic variables (i.e., gender and years of teaching experience) and domains of TSE and JS. As can be seen in Table 1, the three TSE domains correlated positively and moderately as well as two domains of JS. In addition, even though the correlations between TSE and JS domains were also positive, they were relatively small in size.

INSERT TABLE 1 ABOUT HERE

Next, ICC1 values for TSE and JS subdomains with country as the cluster variable were small and ranged from .002 to .005. In contrast, ICC1 values with school as the cluster variable ranged from .059 to .069 (see Table 1) indicating the need to attend to the hierarchical structure of the data. Failure to account for nestedness could lead to biased estimates of standard errors. However, since the measurement invariance testing with an individual-level grouping variable (e.g., teacher gender or career stage) in multilevel data is complicated (i.e., within the same cluster the observations are not independent due to shared group membership), all analyses were conducted under the *Mplus* option TYPE = COMPLEX that successfully corrects miscalculations of standard errors (Kim, Kwok, & Yoon, 2012; Muthén & Satorra, 1995). More specifically, the TYPE = COMPLEX command utilizes a single-level approach in which standard errors of the parameter estimates are adjusted for the complexity of sampling (Asparaouhov, 2005). In support of this approach, previous research indicated that both multilevel CFA and unilevel CFA using the TYPE = COMPLEX option yield equivalent performance when model structures are identical at both levels, which was the case in our research as well (Wu & Kwok, 2012; Muthén & Satorra, 1995).

Measurement Model

We examined the baseline five-factor CFA model and the model fit was good (see Table 2). All standardized factor loadings were statistically significant ($p < .001$) and varied in magnitude from .584 to .839, which indicated the dependability of the indicators to the first-order factors. The intercorrelations among the factors were statistically significant ($p < .001$) and the effect sizes ranged from .17 to .80 (see Table 1). Expectedly, all latent correlations were somewhat greater than their manifest counterparts due to reduction of measurement error in CFA.

INSERT TABLE 1 ABOUT HERE

Measurement and Covariance Invariance

The results of MG-CFA showed that configural invariance models did not have worse fit in comparison to the baseline five-factor model across gender (ΔCFI = .002, ΔRMSEA = .001), career stage (ΔCFI = .003, ΔRMSEA = .001), and educational level (ΔCFI = .001, ΔRMSEA = .000). In addition, metric invariance models did not substantially lose their fit in comparison to their respective configural invariance models across gender (ΔCFI = .001, ΔRMSEA = -.001), career stage (ΔCFI = .001, ΔRMSEA =.-001), and educational level (ΔCFI = .000, ΔRMSEA = -.002) either. These results indicate that sufficient amount of measurement invariance across moderators was achieved, thus allowing the further test of the invariance of the association between domains of TSE and JS.

As metric invariance was established, in the next step we compared the metric invariance models to covariance invariance models in which all latent correlations were constrained to be equal across moderators. The results showed that latent correlations between TSE and JS domains were invariant in their size across gender (ΔCFI = .000, ΔRMSEA = -.001), career stage (ΔCFI = .000, ΔRMSEA =.-001), and educational level (ΔCFI = .000, ΔRMSEA = -.001). Therefore, our findings using TALIS 2018 data indicate that the association between TSE and JS is invariant across gender, career stage, and educational level. In other words, our findings support the assumption of that the association between the two constructs do not vary across teacher factors, which has often been modelled and applied in research and practice (see Gkolia et al., 2014; Zee & Koomen, 2016 for reviews). However, the current data is cross-sectional and does not provide insight into the temporal or causal ordering of the relationship between TSE and JS.

Study 2

Even though the results of the Study 1 based on large TALIS data set clearly indicate the robustness of the association between the TSE and JS across important teacher factors, empirical evidence on the temporal ordering of two constructs is still lacking.

That is, does TSE predict future JS or does JS predict future TSE has remained an answered question. However, determining the temporal ordering of TSE and JS as well as the invariance of their longitudinal relationship is critical to providing an accurate picture of the nature of the constructs' associations with each other, which could have great importance for both theory development and practical implications. Therefore, Study 2 was longitudinal and based on a three-wave full-panel design. In line with the premises of Bandura's (1997) social cognitive theory and Fredrickson's (2004) broaden-and-build theory, but also in line with previous empirical findings, we hypothesize that current TSE levels will predict future JS levels (H2), and current JS levels will predict future TSE levels (H3). Furthermore, we hypothesize that the longitudinal relationship between TSE and JS will be invariant across three moderators – gender, career stage, and educational level (H4), as well as across time (H5).

Method

Participants and Procedure

A convenience sample of 3,002 teachers (82% female) from 135 Croatian state schools participated in a study based on a full panel design with three time points with six months lags (Autumn, 2015; Spring, 2016; Autumn, 2016). At the first measurement occasion, teachers were 41.75 years old (*SD* = 10.44) and had 15.28 (*SD* = 10.50) years of teaching experience. As in Study 1, teachers were grouped into three classes based on their teaching experience (Gu & Day, 2007), and at first measurement occasion. Specifically, 26.71% of the teachers were at early career stage, 47.04% at mid-career stage, and 22.22% at late career stage, while 4.03% of the teachers failed to provide this information. Of the total sample, 28.88% teachers taught at elementary level, 35.18% at middle school level, and 31.15% at secondary school level (others taught either at different levels or did not provide this information).

Data collection procedure was the same on all three occasions and was conducted via postal service and with the assistance of school psychologists. The tasks of school psychologists were the distribution of questionnaires to teachers and returning the questionnaires to the research team upon their completion. Participation in the study was voluntary and anonymous – to preserve teacher anonymity, teachers' responses from the three time points were matched based on specially created codes and the questionnaires were returned to the school psychologists closed each in its own envelope. Of the approached teachers, approximately 50% of them agreed to participate at the first measurement occasion. Of the initial sample of 3,002 teachers, there were 1,525 (51%) teachers at the second wave of data collection and 1,081 (36%) teachers at the third wave of data collection.

To test whether this dropout is related to teacher demographics (i.e., gender, career stage, and educational level) or to the substantive variables (i.e., TSE and JS), an attrition analysis was conducted. The results showed that male teachers were more likely than female teachers to drop out at Time 2 ($\chi^2(1) = 11.36$, $p < .01$) and at Time 3 ($\chi^2(1) = 11.36$, $p < .01$) when compared to ratio of male and female teachers at Time 1. Next, in contrast to elementary and middle-school teachers, greater number of high-school teachers left at Time 3 in comparison to the same number at Time 1 ($\chi^2(2) = 40.49$, $p < .01$) and Time 2 ($\chi^2(2) = 28.13$, $p < .01$). Teachers who left the study right after Time 1 had somewhat lower TSE than those who remained in the study at Time 2 as well, $t(2944) = 2.09$, $p = .037$, $d = .08$. Similarly, teachers who participated in the study both at Time 1 and Time 2 had slightly lower JS than those who dropped out after Time 1, $t(2944) = 2.14$, $p = .046$, $d = .08$. No differences in mean levels of TSE and JS between teachers who participated at all three time points and those who dropped out after Time 2 were found. At last, completers did not differ from non-completers at any time point on years of teaching experience.

The results of attrition analysis indicated that teacher demographic characteristics should be included as covariates into the main analysis. However, since only two mean differences in substantive variables (i.e., TSE and JS) with quite small effect sizes ($d < .20$) emerged, it was decided to proceed with the full information likelihood procedure (FIML; Enders, 2010) to handle the missing data. In support of this decision, the FIML procedure has proved to be an appropriate method to manage the missing data in longitudinal studies (Jeličič et al., 2009).

Measures

To measure TSE, Teacher Self-efficacy Scale (TSES; Schwarzer et al., 1999) was used. Teachers rated their level of agreement for each of the 10 items on a four-point Likert-scale ranging from 1 *(not at all true)* to 4 *(exactly true)*. Sample item is: "When I try really hard, I am able to reach even the most difficult students."

JS was assessed using the Job Satisfaction Scale (Judge et al., 1998), consisting of 5 items measuring overall JS on a five-point Likert-scale ranging from 1 *(completely disagree)* to 5 *(completely agree)*. Sample item is: "I feel fairly satisfied with my present job." Internal consistency coefficients (Cronbach α) of the scales are presented in Table 3.

Data Analyses

Statistical analyses were conducted in several steps. *First*, the Pearson correlation coefficients were calculated to test the associations between TSE and JS between and within assessment occasions, while the ICC1s were calculated to examine whether a substantial amount of variability could be located at the higher unit of analysis (i.e., school level). ICC1 values smaller than .05 would justify running the analyses only at the teacher level (LeBreton & Senter, 2008). *Second*, the measurement invariance (i.e., configural and metric; Millsap,

2001) and covariance invariance of TSE and JS across gender, career stages, and educational levels were tested. Additionally, as a precondition of testing the invariance of structural relationships between latent variables assessed at different time points, their measurement invariance across time was evaluated. As in Study 1, scale items were used in all measurement models as indicators of each of the two latent variables (i.e., TSE and JS) and the residuals of these items were allowed to correlate with each other across time (Marsh & Hau, 1996).

Third, in order to establish the stability and cross-lagged paths between TSE and JS, a series of autoregressive cross-lagged panel models (CLPM) was tested (Campbell, 1963; Kenny & Harackiewicz, 1979). Beyond inherent inclusion of first-order paths (i.e., between adjacent time points) in such models, in order to establish the reciprocal relationship between TSE and JS, it is recommended to consider higher-order paths as well as this allows examination of long-term associations (Arens et al., 2017; Marsh & O'Mara, 2008). Moreover, it was suggested that such analyses should be carried away by starting with a *full-forward model* that includes estimations of all paths; the full-forward model should be then compared to more parsimonious, alternative nested models (Marsh, Byrne, & Yeung, 1999). By following these guidelines, four structural models were hypothesized and tested: a model with both first- and higher-order autoregressive and cross-lagged paths or the full-forward model (M1); a model with first- and higher-order autoregressive paths but only first-order cross-lagged paths (M2); a model with first-order autoregressive paths and first- and higher-order cross-lagged paths (M3); and a model with only first order autoregressive and cross-lagged paths (M4). *Lastly*, the best fitting structural model was tested for the invariance of autoregressive and cross-lagged paths across the four factors in order to examine their robustness.

As in Study 1, all analyses (except descriptive and correlational) were conducted using *Mplus 8.0* (Muthén & Muthén, 2017). The parameters in the models were estimated using the robust maximum-likelihood method (MLR). To evaluate the overall fit of the tested models, we used the same indexes as in Study 1, while measurement invariance of the models was judged based on ΔCFI < .010 and ΔRMSEA < .015 criteria (Chen, 2007; Cheung & Rensvold, 2002). However, in order to uncover the fine-grained differences in fit of structural models, the overall quality of model fit was additionally evaluated based on Akaike information criteria (AIC), whereby an increase of greater than 10 suggests a worse fitting (Burnham & Anderson, 2002).

Results and Discussion

Correlations

As presented in Table 3, teachers with higher levels of TSE also reported higher levels of JS. This relationship was robust between and within different time points. In addition, more experienced teachers reported somewhat higher TSE at Time 2 (r = .07), and higher JS (r = .10) at Time 3. Intraclass correlation coefficients (ICC1) for TSE and JS at different time points were low and ranged from .001 to .023. However, since Bliese (1998) showed that even 1% of the variance (i.e., ICC1=0.01) that is attributed to group membership might result in relatively strong relationship at higher level of analysis, to account for hierarchical nature of the data in Study 2 and to avoid biased estimates of standard errors, the TYPE = COMPLEX option in *Mplus* was again used.

INSERT TABLE 3 ABOUT HERE

Measurement and Covariance Invariance

The results of the test of measurement and covariance invariance across gender, career stage, educational level, and time are presented in Table 4. Imposing restrictions (i.e., equal factor loadings across the four factors) in the metric invariance models and comparing them with the less restrictive configural invariance models, did not result in the loss of model fit across gender (ΔCFI = .000 and ΔRMSEA = -.001), career stage (ΔCFI = .001 and ΔRMSEA = -.001), educational level (ΔCFI = .002 and ΔRMSEA = .000), and time (ΔCFI = .001 and ΔRMSEA = .000). Similar results were obtained after imposing invariant covariances between latent variables – compared to metric invariance models, covariance invariance models did not worsen if fit across gender (ΔCFI = .000 and ΔRMSEA = .00), career stage (ΔCFI = .001 and ΔRMSEA = .000) or educational level (ΔCFI = .000 and ΔRMSEA = .000). In conclusion, necessary amount of measurement and covariance invariance was achieved, which enabled us to proceed with the further tests of structural equivalence.

INSERT TABLE 4 ABOUT HERE

Structural Models

Four hypothesized models with different combinations of first- and higher-order autoregressive and cross-lagged paths were tested and compared to reveal the nature of the association between TSE and JS that is best supported by the data. The results of these analyses are shown in Table 4 as well. When compared to full-forward M1 (both first- and higher-order autoregressive and cross-lagged paths), M2 (first- and higher-order autoregressive paths but only first-order cross-lagged paths) did not show any loss of fit (ΔCFI = .000, ΔRMSEA = .000, ΔAIC = 3.999). Next, in comparison to M1, models M3 (first autoregressive but first- and higher-order cross-lagged paths) and M4 (only first-order autoregressive and cross-lagged paths) did not have substantially worse fit with respect to ΔCFI and ΔRMSEA values (i.e., ΔCFI = .002 and ΔRMSEA = .001 for M3 and ΔCFI = .003 and ΔRMSEA = .001 for M4). However, both models M3 and M4 had worse fit than M1 based on ΔAIC values (ΔAIC = 72.33 and ΔAIC = 105.068, respectively). Similar results were found when comparing models M3 and M4 with M2 — based on ΔAIC values but not ΔCFI and ΔRMSEA values, M3 (ΔCFI = .002, ΔRMSEA = .001, ΔAIC = 68.333), and M4 (ΔCFI = .003, ΔRMSEA = .001, ΔAIC = 100.069) had worse fit to the data. Since both models M1 and M2 had superior fits when contrasted to M3 and M4, and since there were no difference in fit between M1 and M2, the more parsimonious model, that is M2 (includes first- and higher order autoregressive paths and first-order cross-lagged paths), was preferred.

This final M2 model is depicted in Figure 1. TSE and JS correlated with each other across all time points (r = .573, p < .01; r = .490, p < .01; r = .469, p < .01 at Time 1, Time 2, and Time 3, respectively). In addition, all autoregressive path coefficients were statistically significant as expected. More specifically, TSE at Time 1 predicted TSE at Time 2 (β = .587, p < .01) and TSE at Time 3 (β = .376, p < .01), while TSE at Time 2 predicted TSE at Time 3 (β = .354, p < .01). Similar pattern of relationships was found for JS as well: JS at Time 1 predicted JS at Time 2 (β = .817, p < .01) and JS at Time 3 (β = .214, p < .01), while JS at Time 2 predicted JS at Time 3 (β = .613, p < .01). Regarding the cross-lagged associations, an interesting pattern emerged. That is, JS at Time 1 predicted TSE at Time 2 (β = .088, p < .01) and JS at Time 2 predicted TSE at Time 3 (β = .142, p < .01). However, contrary to expectations, TSE failed to predict future levels of JS (β_{T1-T2} = -.063, p > .05 and β_{T2-T3} = .029, p > .05).

INSERT FIGURE 1 ABOUT HERE

Structural Equivalence

In the next step, concurrent latent correlations between TSE and JS, autoregressive paths, and cross-lagged paths specified in model M2 were tested for their equivalence across gender, career stage, educational level, and time (see Table 4). More specifically, M2 model with constrained structural paths across gender, career stage, and educational level was compared to M2 model with unconstrained structural paths (model numbers. In addition, in order to establish the equivalence of structural paths across time, M2 model with paths constrained across time was contrasted to unconstrained M2 model (model number 13 in Table 4). Based on ΔCFI and ΔRMSEA, it can be concluded that the longitudinal structural paths found in M2 are equivalent across gender (ΔCFI = .000 and ΔRMSEA = .000), career stage (ΔCFI = .001 and ΔRMSEA = .000), educational level (ΔCFI = .001 and ΔRMSEA = .000), and time (ΔCFI = .001 and ΔRMSEA = .000).

In sum, our findings using a three-wave longitudinal data indicate that JS predicts TSE rather than the other way around, and that this association is robust across gender, career stage, educational level, and time. The results on the robustness of the association between TSE and JS are in line with Study 1 findings, with the additional factor of time, and with the widely held assumption that findings on TSE–JS associations can be applied regardless of teacher factors. However, our study challenges the widely held assumption that TSE predicts JS.

General Discussion

Examining the invariance of the association between TSE and JS based on two methodologically different but sound studies provides us with nuanced details regarding the extent to which the findings on TSE and JS can be generalized, which can then be helpful for suggesting practical implications to the wider teaching profession. Moreover, ascertaining the order of constructs is critical to understanding which construct must be targeted to increase the outcome of another construct.

Our findings indicate that TSE and JS are positively related to each other concurrently and this association was established on both TALIS and national samples of teachers. Moreover, as our Study 1 suggested, the cross-sectional association between two constructs was indeed invariant across common teachers' factors of gender, career stage, and educational level. These results confirm our first hypothesis (H1) and provide assurance to the proposals that previous studies have made regarding the implications of the positive associations between TSE and JS (see Gkolia et al., 2014; Zee & Koomen, 2016 for reviews). It seems that positive relationship between these constructs that have been oftentimes studied among teachers is indeed robust and stable across different personal and contextual condition.

Even though we hypothesized that TSE would predict future JS (H2) and that JS would predict future TSE (H3), the longitudinal design-based Study 2 points to the unidirectional association between the examined constructs. More specifically, challenging a widely held belief, JS seems to predict future TSE and not the other way around. Our findings contribute to the growing evidence that challenges the assumption that TSE is an antecedent construct (e.g., Author, 2020b; Author, 2020c; Holzberger et al., 2014; Praetorius et al., 2017). It has often been assumed that one's positive experience in their job (e.g., a mastery experience) leads one to feel confident about doing their job, which in turn leads to having a positive evaluation of their job. However, our findings seem to indicate that one's positive experience directly feeds into having a positive evaluation of their job, which in turn leads them to feel confident about doing their job. In fact, this process would be in line with social-cognitive theory (Bandura, 1997), which proposes that affective factors (e.g., satisfaction) can be a source of self-efficacy, as has also been found in the case of teacher burnout (Author, 2020b). Similarly, as suggested by broaden-and-build theory (Fredrickson, 2004), positive affective experiences (i.e., satisfaction) can broaden the repertoire of teachers' cognitive processes and useful instructional tactics and strategies, which, in turn, could enable mastery experience to teachers and positively impact their self-efficacy beliefs.

Furthermore, similar to results from Study 1, findings based on a longitudinal design suggested that temporal association in which JS precedes TSE is stable across teachers' gender, career stage, and educational level, which is consistent with our fourth hypothesis (H4). Moreover, this association seemed invariant across time as well, as suggested by our last hypothesis (H5). Despite different designs (cross-sectional vs. longitudinal) and conceptualizations and operationalizations of TSE and JS (i.e., multi-faceted vs. general), we found consistent findings across the two studies. As such, the findings emphasize the robustness and strength of the examined relationship regardless of the differences in the studies' theoretical or methodological approaches.

Theoretical and Practical Implications

Theoretical models should reconsider their assumptions regarding the direction of association between TSE and JS. Specifically, these models should specify that JS in fact predicts TSE and not the way around. This finding complements the social cognitive model of work and life satisfaction (Lent & Brown, 2006, 2008) and its empirical finding (Lent et al., 2011) that JS predicts life satisfaction in a sample of school teachers. Thus, revising theoretical models that acknowledges that JS can affect both contextualized (e.g., TSE) and non-contextualized experiences (e.g., life satisfaction) may be beneficial.

Some researchers have argued that JS is underlined by a trait-like personality, whereby people with certain characteristics tend to be more satisfied with their jobs than others (see Dormann & Zapf, 2001 for a discussion). The implications of such a claim (and their findings) are contentious — if the dispositional influence on JS is small compared to organizational influences, the content of personnel selection procedures may require a major paradigm shift. However, without further research on the validity of this case in the teaching profession, it is recommended that this is not applied in practice.

A more sensible approach to implementing the study findings is examining ways to enhance teachers' JS. One way to do this is to target its predictors. Working conditions of the teachers, such as the support they receive from school administration, can affect JS (Ma & MacMillan, 1999). As such, ensuring the provision of sufficient administrative support for teachers can be a direct way to enhance their JS. School climate is also important for teachers' JS, including more desirable student behavior and greater student motivation (Collie et al., 2012), and thus investments in trainings for classroom management, for example, may be effective. Moreover, pre-service and in-service professional development opportunities could be provided to develop their ability to deal with stressors at work, such as regulating emotions (Author, 2017). For example, mindfulness training can help them regulate their emotions and increases their JS (Hülsheger et al., 2013).

Limitations and Future Research Directions

Given that TSE and JS variables were both self-reported, the data may have been clouded by subjectivity (Paulhus, 2002). On a related note, assessing all constructs using the same method (i.e., self-report) can lead to common-method bias especially when they are assessed within a single time point (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003). However, since both TSE and JS inherently aim to capture one's subjective experiences and beliefs related to their job, self-report methodology is arguably still be the most appropriate method to collect this type of information. Nevertheless, future studies may wish to complement the self-reported data with other sources of report and mixed methodologies, such as colleagues' report of the target teachers' JS levels, principals' reports of the teachers' efficacy levels in classroom management, and classroom observation ratings. Next, since Study 2 employed convenience sample of teachers, generalizations of its results may be to some extent limited. In particular, even though a ratio between teachers who agreed to participate in the study and all the teachers who were contacted is surprisingly high (50%), characteristics of teachers who declined to enroll have remained unknown. Future studies should aim at targeting more representative samples of teachers at both national and international level in order to obtain results that could be more easily generalizable to whole teaching population.

Underlying our study is the belief that, in line with other studies and models (e.g., Klassen & Chiu, 2010; Schleicher, 2018), TSE and JS should be increased in order to address the teacher shortage crisis. However, such a link has rarely been empirically tested. Therefore, future studies should endeavor to collect such data to interrogate the extent to which interventions targeting JS and TSE can help address this international problem. Finally, our studies examined the direct association between TSE and JS. It is possible, however, that there may be mediators or moderators of the association. For example, the model of job stress and satisfaction in teachers (Troesch & Bauer, 2017), proposes that teachers' positive and negative appraisals of the job is a mediator between TSE and JS. For the benefit particularly of theory development, studies on the mediators and moderators of the associations are recommended.

Conclusion

In order to most effectively strengthen the teaching profession, we must learn more their experiences and address the needs that they have in the ways that we can. Our two studies have found that JS is imperative to how efficacious a teacher feels about doing their job, irrespective of their gender, their years of teaching experience, and the educational level they teach in. As such, researchers, practitioners, and policymakers are encouraged to work together to provide teachers the appropriate resources, environments, and training that can help them stay and continue to do what they enjoy the most — teach.

References

Author (2020a)

Author (2020b)

Author (2020c)

Author (2018)

Author (2017)

Adebomi, O., & Olufunke, I. H. (2012). Job satisfaction and self-efficacy as correlates of job commitment of special education teachers in Oyo State. *Journal of Education and Practice*, *3*(9), 95–103.

Ainley, J., & Carstens, R. (2018). *Teaching and Learning International Survey (TALIS) 2018 Conceptual Framework* (No. 187; OECD Education Working Papers). https://doi.org/10.1787/799337c2-en

Akomolafe, M. J., & Ogunmakin, A. O. (2014). Job satisfaction among secondary school teachers: Emotional intelligence, occupational stress and self-efficacy as predictors. *Journal of Educational and Social Research*, *4*(3), 487–498.

Asparouhov, T. (2006). General multi-level modeling with sampling weights. *Communications in Statistics—Theory and Methods, 35*(3), 439-460.

Avanzi, L., Miglioretti, M., Velasco, V., Balducci, C., Vecchio, L., Fraccaroli, F., & Skaalvik, E. M. (2013). Cross-validation of the Norwegian Teacher's Self-Efficacy Scale (NTSES). *Teaching and Teacher Education*, *31*, 69–78.

Bandura, A. (1986). *Social foundations of thought and action*. Prentice-Hall.

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

Bliese, P. D. (1998). Group size, ICC values, and group-level correlations: A simulation. *Organizational research methods, 1*(4), 355-373.

Brown, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A.

Bollen, & J. S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Sage.

Buchanan, J., Prescott, A., Schuck, S., Aubusson, P., Burke, P., & Louviere, J. (2013). Teacher retention and attrition: Views of early career teachers. *Australian Journal of Teacher Education*, *38*(3), 112–129.

Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multimodel inference: A practical information-theoretic approach* (2nd ed.). Springer.

Campbell, D. T. (1963a). From description to experimentation: Interpreting trends as quasiexperiments. In: C. W. Harris (ed.), *Problems in Measuring Change*. Madison, WI: University of Chicago Press, pp. 212–242.

Canrinus, E. T., Helms-Lorenz, M., Beijaard, D., Buitink, J., & Hofman, A. (2012). Self-efficacy, job satisfaction, motivation and commitment: Exploring the relationships between indicators of teachers' professional identity. *European Journal of Psychology of Education*, *27*(1), 115–132.

Caprara, G. V., Barbaranelli, C., Steca, P., & Malone, P. S. (2006). Teachers' self-efficacy beliefs as determinants of job satisfaction and students' academic achievement: A study at the school level. *Journal of School Psychology*, *44*(6), 473–490.

Carson, R. L., Baumgartner, J. J., Ota, C. L., Kuhn, A. P., & Durr, A. (2017). An ecological momentary assessment of burnout, rejuvenation strategies, job satisfaction, and quitting intentions in childcare teachers. *Early Childhood Education Journal*, *45*(6), 801–808.

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.

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.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, *9*(2), 233–255.

Collie, R. J., Shapka, J. D., & Perry, N. E. (2012). School climate and social–emotional learning: Predicting teacher stress, job satisfaction, and teaching efficacy. *Journal of Educational Psychology*, *104*(4), 1189–1204.

Crossman, A., & Harris, P. (2006). Job satisfaction of secondary school teachers. *Educational Management Administration & Leadership*, *34*(1), 29–46.

Department for Education. (2019). *Teacher recruitment and retention strategy*. https://www.gov.uk/government/publications/teacher-recruitment-and-retention-strategy

Dimitrov, D. M. (2010). Testing for factorial invariance in the context of construct validation. *Measurement and Evaluation in Counseling and Development*, *43*(2), 121–149.

Dormann, C., & Zapf, D. (2001). Job satisfaction: A meta-analysis of stabilities. *Journal of Organizational Behavior*, *22*(5), 483–504.

Dupriez, V., Delvaux, B., & Lothaire, S. (2016). Teacher shortage and attrition: Why do they leave? *British Educational Research Journal*, *42*(1), 21–39.

Education for All Global Monitoring Report and the UNESCO Education Sector. (2015). *The challenge of teacher shortage and quality: Have we succeeded in getting enough quality teachers into classrooms?* (Policy Paper 19). UNESCO. http://unesdoc.unesco.org/images/0023/002327/232721e.pdf

Enders, C. K. (2010). *Applied Missing Data Analysis*. Guilford Press.

Federici, R. A., & Skaalvik, E. M. (2012). Principal self-efficacy: Relations with burnout, job satisfaction and motivation to quit. *Social Psychology of Education: An International Journal*, *15*(3), 295–320.

Fredrickson, B. L. (2004). The broaden–and–build theory of positive emotions. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 359(1449), 1367-1377.

Gist, M. E., & Mitchell, T. R. (1992). Self-efficacy: A theoretical analysis of its determinants and malleability. *Academy of Management review, 17*(2), 183-211.

Gkolia, A., Belias, D., & Koustelios, A. (2014). Teacher's job satisfaction and selfefficacy: A review. *European Scientific Journal*, *10*(22), 321–342.

Gu, Q., & Day, C. (2007). Teachers resilience: A necessary condition for effectiveness. *Teaching and Teacher Education*, *23*(8), 1302–1316.

Hamaker, E. L. (2005). Conditions for the equivalence of the autoregressive latent trajectory model and a latent growth curve model with autoregressive disturbances. *Sociological Methods & Research*, *33*(3), 404–416.

Holzberger, D., Philipp, A., & Kunter, M. (2014). Predicting teachers' instructional behaviors: The interplay between self-efficacy and intrinsic needs. *Contemporary Educational Psychology*, *39*(2), 100–111.

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.

Hülsheger, U. R., Alberts, H. J. E. M., Feinholdt, A., & Lang, J. W. B. (2013). Benefits of mindfulness at work: The role of mindfulness in emotion regulation, emotional exhaustion, and job satisfaction. *Journal of Applied Psychology*, *98*(2), 310–325.

IBF International Consulting. (2013). *Study on policy measures to improve the attractiveness of the teaching profession in Europe: Final report. Volume 2*. Publications Office of the European Union.

Ingersoll, R. (2003). *Is there really a teacher shortage?* University of Washington: Center for the Study of Teaching and Policy.

Jeličič, H., Phelps, E., & Lerner, R. M. (2009). Use of missing data methods in longitudinal studies: the persistence of bad practices in developmental psychology. *Developmental Psychology*, *45*(4), 1195–1199.

Judge, T. A., Locke, E. A., Durham, C. C., & Kluger, A. N. (1998). Dispositional effects on job and life satisfaction: The role of core evaluations. *Journal of Applied Psychology*, *83*(1), 17–34.

Judge, T. A., Thoresen, C. J., Bono, J. E., & Patton, G. K. (2001). The job satisfaction--job performance relationship: A qualitative and quantitative review. *Psychological Bulletin*, *127*, 376–407.

Kenny, D. A., & Harackiewicz, J. M. (1979). Cross-lagged panel correlation: Practice and promise. *Journal of Applied Psychology*, *64*(4), 372–379.

Kim, E. S., Kwok, O. M., & Yoon, M. (2012). Testing factorial invariance in multilevel data: A Monte Carlo study. *Structural Equation Modeling: A Multidisciplinary Journal, 19*(2), 250-267.

LeBreton, J. M., & Senter, J. L. (2008). Answers to 20 questions about interrater reliability and interrater agreement. *Organizational Research Methods, 11*(4), 815-852.

Klassen, R. M., & Chiu, M. M. (2010). Effects on teachers' self-efficacy and job satisfaction: Teacher gender, years of experience, and job stress. *Journal of Educational Psychology*, *102*(3), 741–756.

Kunter, M., Klusmann, U., Baumert, J., Richter, D., Voss, T., & Hachfeld, A. (2013). Professional competence of teachers: Effects on instructional quality and student development. *Journal of Educational Psychology*, *105*(3), 805–820.

Lent, R. W., & Brown, S. D. (2006). On conceptualizing and assessing social cognitive constructs in career research: A measurement guide. *Journal of Career Assessment*, *14*(1), 12–35.

Lent, R. W., & Brown, S. D. (2008). Social cognitive career theory and subjective well-being in the context of work. *Journal of Career Assessment*, *16*(1), 6–21.

Lent, R. W., Nota, L., Soresi, S., Ginevra, M. C., Duffy, R. D., & Brown, S. D. (2011). Predicting the job and life satisfaction of Italian teachers: Test of a social cognitive model. *Journal of Vocational Behavior*, *79*(1), 91–97.

Lindqvist, P., Nordänger, U. K., & Carlsson, R. (2014). Teacher attrition the first five years--A multifaceted image. *Teaching and Teacher Education*, *40*, 94–103.

Locke, E. A. (1969). What is job satisfaction? *Organizational Behavior and Human Performance*, *4*(4), 309–336.

Malinen, O.-P., & Savolainen, H. (2016). The effect of perceived school climate and teacher efficacy in behavior management on job satisfaction and burnout: A longitudinal study. *Teaching and Teacher Education*, *60*, 144–152.

Marsh, H. W., Byrne, B. M., & Yeung, A. S. (1999). Causal ordering of academic self-concept and achievement: Reanalysis of a pioneering study and. *Educational Psychologist, 34*(3), 155-167.

Marsh, H. W., & Hau, K.-T. (1996). Assessing goodness of fit: Is parsimony always desirable? *Journal of Experimental Education*, *64*(4), 364–390.

Ma, X., & MacMillan, R. B. (1999). Influences of workplace conditions on teachers' job satisfaction. *The Journal of Educational Research*, *93*(1), 39–47.

Millsap, R. E. (2001). When trivial constraints are not trivial: The choice of uniqueness constraints in confirmatory factor analysis. *Structural Equation Modeling, 8*(1), 1-17.

Mottet, T. P., Beebe, S. A., Raffeld, P. C., & Medlock, A. L. (2004). The effects of student verbal and nonverbal responsiveness on teacher self-efficacy and job satisfaction. *Communication Education*, *53*(2), 150–163.

Muthen, B. O., & Satorra, A. (1995). Complex sample data in structural equation modeling. *Sociological Methodology*, 267-316.

Muthén, L. K., & Muthén, B. O. (2017). *Mplus Version 8 User's Guide* (Eighth). Muthén & Muthén.

NSW Government. (2020, May 7). *Teacher subsidies, allowances and bonuses*. https://education.nsw.gov.au/about-us/careers-at-education/salary-and-benefits/subsidies-allowances-and-bonuses

OECD. (2019a). *TALIS 2018 and TALIS Starting Strong 2018 User Guide*. http://www.oecd.org/education/talis/TALIS_2018-TALIS_Starting_Strong_2018_User_Guide.pdf

OECD. (2019b). *TALIS 2018 Technical Report*. https://www.oecd.org/education/talis/TALIS_2018_Technical_Report.pdf

OECD. (2019c). *PISA 2018 Results (Volume II) Where All Students Can Succeed: Where All Students Can Succeed*. OECD Publishing.

OECD. (2020). *TALIS 2018 Results (Volume II) Teachers and School Leaders as Valued Professionals*. https://www.oecd-ilibrary.org/education/annex-bmain-breakdown-variables_d1ba43b3-en

Organ, D. W., & Ryan, K. (1995). A meta-analytic review of attitudinal and dispositional predictors of organizational citizenship behavior. *Personnel Psychology*, *48*(4), 775–802.

Paulhus, D. L. (2002). Socially desirable responding: The evolution of a construct. In H. Braun, D. N. Jackson, & D. E. Wiley (Eds.), *The role of constructs in psychological and educational measurement* (pp. 49–69). Erlbaum.

Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology, 88*(5), 879-903.

Praetorius, A.-K., Lauermann, F., Klassen, R. M., Dickhäuser, O., Janke, S., & Dresel, M. (2017). Longitudinal relations between teaching-related motivations and student-reported teaching quality. *Teaching and Teacher Education*, *65*, 241–254.

Richardson, P. W., Karabenick, S. A., & Watt, H. M. G. (2014). *Teacher Motivation: Theory and Practice*. Routledge.

Rudnev, M., Lytkina, E., Davidov, E., Schmidt, P., & Zick, A. (2018). Testing measurement invariance for a second-order factor. A cross-national test of the alienation scale. *Methods, Data, Analyses*, *12*(1), 47-76.

Salanova, M., Llorens, S., & Schaufeli, W. B. (2011). "Yes, I can, I feel good, and I just do it!" On gain cycles and spirals of efficacy beliefs, affect, and engagement. *Applied Psychology*, *60*(2), 255–285.

Schleicher, A. (2018). *Valuing our Teachers and Raising their Status: How communities can help* (International Summit on the Teaching Profession,). OECD Publishing. https://doi.org/10.1787/23127090

Schwarzer, R., Schmitz, G. S., & Daytner, G. T. (1999). *The teacher self-efficacy scale*. http://userpage.fu-berlin.de/gesund/skalen/Language_Selection/Turkish/Teacher_Self-Efficacy/teacher_self-efficacy.htm

Skaalvik, E. M., & Skaalvik, S. (2007). Dimensions of teacher self-efficacy and relations with strain factors, perceived collective teacher efficacy, and teacher burnout. *Journal of Educational Psychology*, *99*(3), 611–625.

Skaalvik, E. M., & Skaalvik, S. (2017). Motivated for teaching? Associations with school goal structure, teacher self-efficacy, job satisfaction and emotional exhaustion. *Teaching and Teacher Education*, *67*, 152–160.

Soto, M., & Rojas, O. (2019). Self-efficacy and job satisfaction as antecedents of citizenship behaviour in private schools. *International Journal of Management in Education*, *13*(1), 82-96.

Taylor, D. L., & Tashakkori, A. (1995). Decision participation and school climate as predictors of job satisfaction and teachers' sense of efficacy. *Journal of Experimental Education*, *63*(3), 217–230.

Troesch, L. M., & Bauer, C. E. (2017). Second career teachers: Job satisfaction, job stress, and the role of self-efficacy. *Teaching and Teacher Education*, *67*, 389–398.

Tschannen-Moran, M., & Hoy, A. W. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and Teacher Education, 17*(7), 783-805.

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

UNESCO Institute for Statistics. (2016). *The world needs almost 69 million new teachers to reach the 2030 education goals* (No. 39). http://uis.unesco.org/sites/default/files/documents/fs39-the-world-needs-almost-69-million-new-teachers-to-reach-the-2030-education-goals-2016-en.pdf

UNESCO Institute for Statistics. (2012). International standard classification of education: ISCED 2011. Montreal: UNESCO Institute for Statistics.

Veldman, I., van Tartwijk, J., Brekelmans, M., & Wubbels, T. (2013). Job satisfaction and teacher–student relationships across the teaching career: Four case studies. *Teaching and Teacher Education*, *32*, 55–65.

Viel-Ruma, K., Houchins, D., Jolivette, K., & Benson, G. (2010). Efficacy beliefs of special educators: The relationships among collective efficacy, teacher self-efficacy, and job satisfaction. *Teacher Education and Special Education*, *33*(3), 225–233.

Wang, H., Hall, N. C., & Rahimi, S. (2015). Self-efficacy and causal attributions in teachers: Effects on burnout, job satisfaction, illness, and quitting intentions. *Teaching and Teacher Education*, *47*, 120–130.

Wu, J. Y., & Kwok, O. M. (2012). Using SEM to analyze complex survey data: A comparison between design-based single-level and model-based multilevel approaches.

Structural Equation Modeling: A Multidisciplinary Journal, 19(1), 16-35.

Zee, M., & Koomen, H. M. Y. (2016). Teacher self-efficacy and its effects on classroom processes, student academic adjustment, and teacher well-being: A synthesis of 40 years of research. *Review of Educational Research*, *86*(4), 981–1015.

Zembylas, M. (2003). Interrogating "teacher identity": Emotion, resistance, and self-formation. *Educational Theory*, *53*(1), 107–127.

Table 1

Descriptive Statistics and Pearson Correlations for Demographics, Self-Efficacy, and Job Satisfaction in Study 1

	Variable	*ICC1*	*M*	SD	Skewness (SE)	Kurtosis (SE)	2	3	4	5	6	7
1	Gender	n/a	n/a	n/a	n/a	n/a	-.02	-.04	-.07	-.03	-.02	-.03
2	Experience (years)	n/a	16.17	10.33	0.50 (0.01)	-0.49 (0.01)	-	.07	.04	.04	-06	.01
3	Self-efficacy in classroom management	.061	3.33	0.56	-0.54 (0.01)	-.24 (0.01)	-	-	.62	.67	.17	.14
4	Self-efficacy in instruction	.059	3.29	0.54	-0.41 (0.01)	-.32 (0.01)	-	.73	-	.69	.17	.14
5	Self-efficacy in student engagement	.065	3.20	0.60	-0.34 (0.01)	-.61 (0.01)	-	.76	.80	-	.19	.16
6	Job satisfaction with work environment	.069	3.12	0.58	-0.50 (0.01)	.41 (0.01)	-	.20	.21	.23	-	.51
7	Job satisfaction with profession	.068	3.04	0.64	-0.46 (0.01)	-.10 (0.01)	-	.17	.18	.20	.59	-

Note. SE= Standard Error; Sums of subscales' scores were calculated based on raw TALIS data; All correlations are statistically significant at $p < .01$; Manifest correlations are presented above the diagonal and latent correlations are presented below the diagonal.

Table 2

Fit Statistics of All Tested Models in Study 1

Model Number	Model Type	χ2 (*df*)	CFI	TLI	RMSEA (90% C.I.)	SRMR	AIC
Measurement Invariance Models							
1	Baseline five-factor model	132190.793 (160)	.925	.911	.058 (.058, .058)	.047	8624313.597
2	Configural invariance across gender	136792.475 (320)	.923	.908	.059 (.058, .059)	.048	8616495.679
3	Metric invariance across gender	137703.645 (335)	.922	.912	.058 (.057, .058)	.048	8617112.401
4	Configural invariance across career stage	138813.178 (480)	.922	.907	.059 (.059, .060)	.048	8576216.206
5	Metric invariance across career stage	139804.032 (510)	.921	.912	.058 (.057, .058)	.049	8576873.541
6	Configural invariance across educational level	124208.918 (480)	.926	.912	.058 (.058, .058)	.047	7998408.909
7	Metric invariance across educational level	124416.629 (510)	.926	.917	.056 (.056, .057)	.048	7999552.236
Covariance Invariance Models							
8	Covariance invariance across gender	138385.269 (345)	.922	.914	.057 (.057, .057)	.051	8617486.142
9	Covariance invariance across career stage	140687.717 (530)	.921	.915	.057 (.057, .057)	.051	8577358.935
10	Covariance invariance across educational level	123974.208 (530)	.926	.921	.055 (.055, .055)	.052	8000168.803

Note: CFI = Comparative Fit Index; TLI = Tucker Lewis Index; RMSEA = root mean square error of approximation; SRMR = Standardized Root Mean Square Residual.

Table 3

Descriptive Statistics, Cronbach Alphas, and Pearson Correlations for Demographics, Self-Efficacy, and Job Satisfaction in Study 2

	Variable	*ICC1*	*M*	*SD*	Skewness (SE)	Kurtosis (SE)	2	3	4	5	6	7	8
1	Gender	n/a	n/a	n/a	n/a	n/a	.05*	.02	.02	-.01	.03	.03	.01
2	Experience (years)	n/a	15.28	10.50	0.50 (.045)	-0.68 (.090)	(n/a)	.03	.07**	.03	.04	.03	.10**
3	Self-efficacy T1	.001	3.37	0.40	-0.56 (.045)	0.79 (.090)		(.84)	.57**	.62**	.48**	.34**	.38**
4	Self-efficacy T2	.011	3.33	0.41	-0.39 (.063)	0.43 (.126)			(.86)	.61**	.38**	.47**	.37**
5	Self-efficacy T3	.006	3.29	0.44	-0.31 (.076)	0.14 (.151)				(.88)	.41**	.40**	.53**
6	Job satisfaction T1	.022	4.12	0.60	-0.87 (.045)	1.34 (.090)					(.83)	.69**	.63**
7	Job satisfaction T2	.023	4.04	0.60	-0.78 (.063)	0.89 (.126)						(.83)	.70**
8	Job satisfaction T3	.016	4.03	0.65	-0.89 (.075)	1.39 (.150)							(.85)

Note. SE= Standard Error. Cronbach αs are shown in parentheses in the diagonal. ** $p < .01$; * $p < .05$.

Table 4

Fit Statistics of All Tested Models in Study 2

Model number	Model type	χ^2 (*df*)	CFI	TLI	RMSEA (90% C.I.)	SRMR	AIC
Measurement /Covariance Invariance Models							
1	Configural invariance across gender	3888.80 (1758)	.933	.925	.029 (.027, .030)	.048	134686.751
2	Metric invariance across gender	3935.85 (1803)	.933	.927	.028 (.027, .030)	.054	134657.510
3	Covariance invariance across gender	3952.83 (1818)	.933	.927	.028 (.027, .029)	.055	134651.675
4	Configural invariance across career stages	4710.35 (2637)	.934	.925	.029 (.027, .030)	.052	130524.421
5	Metric invariance across career stages	4834.31 (2727)	.933	.926	.028 (.027, .030)	.068	130490.687
6	Covariance invariance across career stages	4885.90 (2757)	.932	.927	.028 (.027, .030)	.068	130498.550
7	Configural invariance across educational levels	4773.51 (2637)	.931	.922	.029 (.028, .031)	.052	129444.892
8	Metric invariance across educational levels	4904.60 (2727)	.929	.925	.029 (.028, .030)	.064	129416.096
9	Covariance invariance across educational levels	4956.69 (2757)	.929	.923	.029 (.028, .030)	.066	129424.760
10	Configural invariance over time	2589.39 (879)	.944	.936	.025 (.024, .027)	.038	136222.785
11	Metric invariance over time	2635.76 (905)	.943	.938	.025 (.024, .026)	.041	136220.357
Higher-order Autoregressive and Cross-Lagged Path Models							
12	M1 first- and higher-order autoregressive and cross-lagged paths (*full-forward*)	2635.76 (905)	.943	.938	.025 (.024, .026)	.041	136220.357
13	M2 first- and higher-order autoregressive paths and first order cross-lagged paths	2643.40 (907)	.943	.937	.025 (.024, .026)	.042	136224.356
14	M3 first order autoregressive paths and first- and higher order cross-lagged paths	2705.27 (907)	.941	.935	.026 (.025, .027)	.044	136292.689
15	M4 only first order autoregressive and cross-lagged paths	2733.96 (909)	.940	.934	.026 (.025, .027)	.049	136325.425
Structural Invariance Models							
16	M2 with structural paths unconstrained across gender	4107.36 (1809)	.928	.921	.029 (.028, .031)	.056	134831.516
17	M2 with structural paths constrained across gender	4128.37 (1819)	.928	.921	.029 (.028, .031)	.060	134839.546
18	M2 with structural paths unconstrained across career stages	5029.03 (2736)	.927	.920	.030 (.028, .031)	.067	130692.377
19	M2 with structural paths constrained across career stages	5062.93 (2756)	.926	.920	.030 (.028, .031)	.071	130690.161
20	M2 with structural paths unconstrained across educational levels	5088.12 (2736)	.924	.917	.030 (.029, .031)	.067	129606.239
21	M2 with structural paths constrained across educational levels	5117.63 (2756)	.923	.917	.030 (.029, .031)	.069	129601.064
22	M2 with structural paths constrained over time	2670.68 (911)	.942	.937	.025 (.024, .027)	.044	136250.061

Note: CFI = Comparative Fit Index; TLI = Tucker Lewis Index; RMSEA = root mean square error of approximation; SRMR = Standardized Root Mean Square Residual; AIC = Akaike Information Criterion.

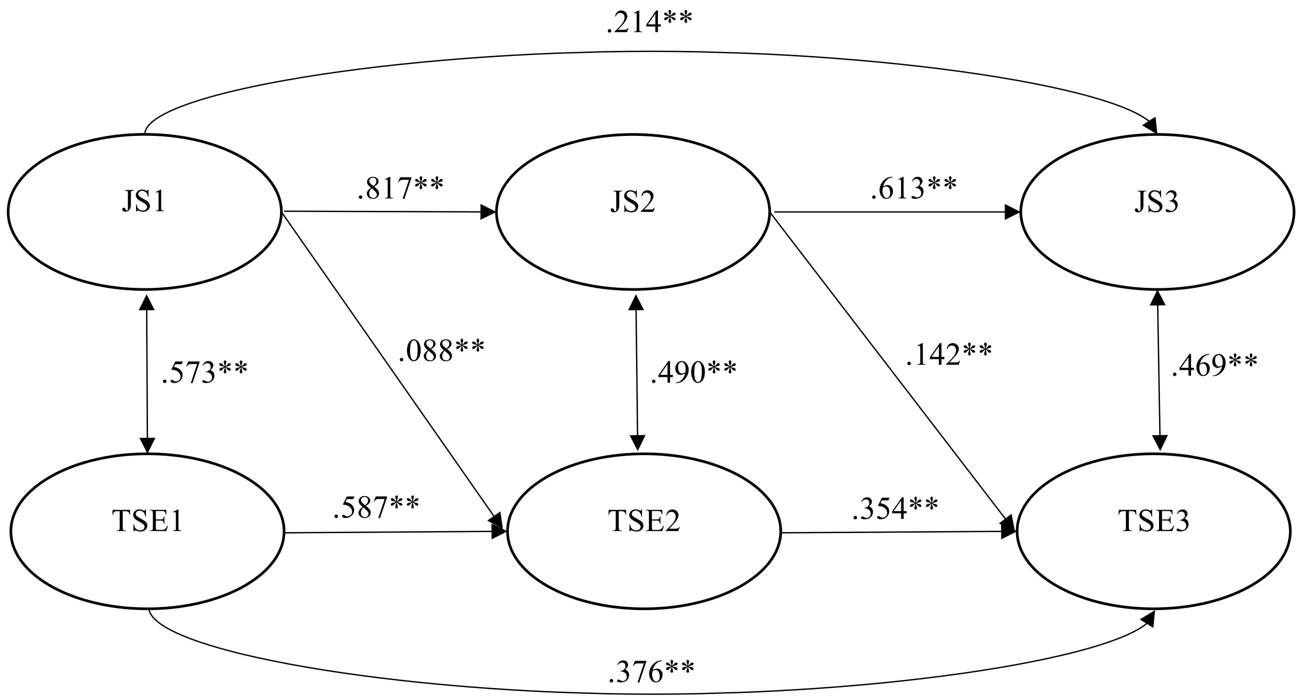

Figure 1. Final M2 model

Note. ***p*<.01; only statistically significant paths are shown. TSE = Teacher self-efficacy; JS = Job satisfaction.