Teacher self-efficacy and burnout: Determining the directions of prediction through an autoregressive cross-lagged panel model

Lisa E. Kim[a] & Irena Burić[b]

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

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

Corresponding author: Correspondence concerning this article should be directed to Lisa Kim, Department of Education, University of York, UK.

E-mail: lisa.kim@york.ac.uk. Phone: +44 1904 322605

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

Date of acceptance: 7 September 2019

Abstract

It is often assumed that low levels of teacher self-efficacy (TSE) leads to negative outcomes, including burnout; however, the temporal order of the construct predictions has rarely been examined. We used an autoregressive cross-lagged panel design to examine whether TSE and burnout are concurrently associated with each other, whether TSE predicts future burnout levels, and/or whether burnout predicts future TSE levels. An initial sample of 3002 Croatian teachers (82% female) from across three educational levels (i.e., elementary, middle, and secondary schools) with varying years of teaching experiences (M=15.28, SD=10.50) completed questionnaires on their levels of TSE and burnout (exhaustion and disengagement) at three time points (at approximately six-month intervals). We found that burnout has a more prominent role in predicting future levels of TSE than TSE does in predicting future levels of burnout. These findings challenge the theoretical and empirical conceptualizations assuming that TSE is a predictor of burnout. Policies and interventions that focus on decreasing teacher burnout rather than increasing TSE levels may be best.

Keywords: teacher burnout; self-efficacy; teacher effectiveness; teacher retention; structural equation modelling

Educational Impact and Implications Statement

To assist in teacher retention and development, policies and interventions have often focused on increasing teacher self-efficacy (TSE), under the untested assumption that low TSE is the root cause of negative outcomes, such as burnout. We found in our analyses, using an autoregressive cross-lagged panel design consisting of an initial sample of 3002 Croatian teachers (82% female), that burnout has a more prominent role in predicting future levels of TSE than TSE does in predicting future levels of burnout. That is, researchers, practitioners, and policymakers may find focusing on decreasing teacher burnout more beneficial than on increasing TSE levels.

Teacher self-efficacy and burnout: Determining the directions of prediction through an autoregressive cross-lagged panel model

Many countries are experiencing a teacher shortage crisis (e.g., Buchanan et al., 2013; Sutcher, Darling-Hammond, & Carver-Thomas, 2016; Vlahović-Štetić & Vizek Vidović, 2005). To tackle this problem, large international agencies have been developing policy guidelines to assist in attracting, developing, and retaining teachers (e.g., Education for All Global Monitoring Report and the UNESCO Education Sector, 2015; Organization for Economic Cooperation and Development, 2005). Recruiting more teachers is a commonly-used strategy to address the teacher shortage crisis (UNESCO Institute for Statistics 2016); however, a shortage will persist as long as there continues to be high rates of teacher attrition (Ingersoll, 2002).

Teachers' intention to quit and their subsequent attrition are often believed to be consequences of teachers experiencing burnout (Chang, 2009) and experiencing low self-efficacy (Wang, Hall, & Rahimi, 2015). In this light, various strategies have been proposed targeting factors associated with burnout and self-efficacy (Tschannen-Moran & Hoy, 2007; Vandenberghe & Huberman, 1999). However, although these two constructs have often been examined together, the nature and the directionality of their association have not been established. That is, it is unknown whether low teacher self-efficacy (TSE) causes burnout and/or burnout causes low TSE, although the former association is often assumed to be true. Understanding which construct is the antecedent and which is the consequence is important for effective policy development and intervention implementation. As such, this study examines the nature and the directionality of the associations between TSE and burnout using three-wave longitudinal data.

Burnout

The teaching profession is perceived to be one of the most stressful professions (Johnson et al., 2005), involving numerous tasks (e.g., class preparation and classroom management) and interactions with multiple groups of people (e.g., students, colleagues, and parents; Jensen, Sandoval-Hernández, Knoll, & Gonzalez, 2012). One indicator of its stressfulness is the high attrition rate; some researchers have quoted as high as 40-50% in the first five years of teaching (Darling-Hammond, 2010; Ingersoll, 2003). Indeed, unmanaged prolonged exposure to occupational stress can lead to burnout (Maslach, Schaufeli, & Leiter, 2001).

Burnout is a multidimensional construct defined by the dimensions of exhaustion (physically, affectively, and cognitively) and disengagement from work and people (Demerouti, Bakker, Vardakou, & Kantas, 2003). Although Maslach and colleagues (2001) specified diminished personal accomplishment as the third dimension of burnout, various theoretical and empirical studies have challenged the validity of this dimension (see Demerouti & Bakker, 2008 for a review). Hence, we consider two dimensions of the burnout construct (i.e., exhaustion and disengagement) and recognize the construct's multidimensionality by modelling these two dimensions in separate analyses in this study.

Experiences of the symptoms of burnout can have a plethora of negative effects on teachers, students, and schools. For example, one can experience mental health difficulties, such as low self-confidence, low self-esteem, and clinical depression (Schonfeld, 2001) and display diminished abilities to tolerate student misconduct, which can potentially magnify student behavioral problems (Huberman, 1993; Lamude & Scudder, 1992). Furthermore, student academic achievement can be negatively affected (Arens & Morin, 2016; Voss, Wagner, Klusmann, Trautwein, & Kunter, 2017). At the school level, teachers experiencing burnout symptoms can be more frequently absent from work (Schonfeld, 2001), which necessitates finding substitute teachers that has administrative and financial implications for the school. Furthermore, some researchers have claimed that burnout symptoms can be contagious— it can spread to colleagues (Bakker & Schaufeli, 2000) and negatively affect general staff morale (Leithwood, Menzies, Jantzi, & Leithwood, 1999). The seriousness of these effects is bolstered by findings that burnout levels are relatively stable over time (Hakanen, Schaufeli, & Ahola, 2008; Pas et al., 2012; Schaufeli, Maasen, Bakker, & Sixma, 2011). Thus, it is important to investigate the ways that teachers' experiences of exhaustion and disengagement can be prevented and/or ameliorated by studying its associations with other factors and constructs.

Teacher Self-Efficacy

A construct that has often been studied in this light is teacher self-efficacy (TSE; Zee & Koomen, 2016). TSE, grounded within the social cognitive theory (Bandura, 1977), is a teacher's belief of how well they are capable of conducting profession-related activities, such as managing the classroom and using instructional strategies (Schwarzer, Schmitz, & Daytner, 1999; Tschannen-Moran & Hoy, 2001). TSE levels are affected by various teacher, classroom, school, and leadership factors. For example, Fackler and Mamberg (2016) examined 14 OECD countries in the Teaching and Learning International Survey (TALIS) dataset to assess the effect of these factors. Fackler and Mamberg found that teachers had higher levels of self-efficacy when they taught students with higher records of academic achievement, the teacher themselves had higher levels of socioeconomic status, were working under principals with greater years of work experience, and were working under principals with a greater sense of instructional leadership style.

Zee and Koomen (2016) reviewed previous studies on the associations TSE has with a variety of classroom processes, student, and teacher outcomes. Overall, they found that students taught by highly self-efficacious teachers were more academically successful, were more motivated, and had more positive attitude towards learning, school satisfaction, and achievement confidence across educational levels. Furthermore, these teachers were more satisfied with their jobs and experienced less job stress, which are said to mediate the effects some factors have in contributing to burnout (Grayson & Alvarez, 2008; Schwarzer & Hallum, 2008). As with burnout, TSE levels seem to be relatively stable over time (e.g., Holzberger et al., 2013; Pas et al., 2012).

Teacher Self-Efficacy and Burnout

The negative association between self-efficacy and job burnout is a well-established finding (Aloe, Amo, & Shanahan, 2014; Skaalvik & Skaalvik, 2010). For example, a meta-analysis reported that across occupations, self-efficacy was negatively associated with job burnout with an average effect size of -.33 (Shoji et al., 2016). The study also reported that the association was stronger for teachers (-.38) than for health-care providers (-.26). In fact, Zee and Koomen (2016) found in their review of self-efficacy levels in teachers that TSE has been consistently negatively associated with burnout, with effect sizes ranging from -.17 to -.63.

Although some researchers have modelled self-efficacy and the dimensions of burnout to be concurrently associated constructs (e.g., Skaalvik & Skaalvik, 2010), other researchers have attributed low TSE levels as the root cause of burnout (Cherniss, 2017; Leiter, 1992). Such conceptualization is driven by the assumption that teacher motivation variables are antecedents of occupational well-being and effective teaching practices (e.g., Kunter et al., 2013; Richardson, Karabenick, & Watt, 2014), especially given that they are relatively stable constructs (Praetorius et al., 2017). Empirical studies using mediational analyses exemplify this assumption. Studies have investigated the link between TSE and burnout using mediators such as job stress (Schwarzer & Hallum, 2008), instructional practices and student stressors in the classroom (Martin, Sass, & Schmitt, 2012), and difficulties related to the classroom (i.e., student diversity and misbehavior; Betoret, 2009). Although researchers have often assumed a unidirectional association (i.e., from TSE to burnout dimensions), the possibility of a different direction of prediction has rarely been examined. One reason for this absence may be due to the difficulty of obtaining longitudinal data and in using advanced statistical techniques to thoroughly test and compare multiple models with different permutations of the predictive paths.

Hobfoll's (1989) Conservation of Resources Theory suggests that the association between self-efficacy and dimensions of burnout may not necessarily be unidirectional. According to this theory, individuals strive to obtain, conserve, and build resources for their positive well-being. Depleted resources result in stress that can manifest physically, emotionally, and/or psychologically. In this light, the prolonged exposure to stressors without resource replacement can result in low levels of self-efficacy and low levels of motivation and commitment to the job. Such a state can then lead to one being emotionally exhausted and detaching themselves from work. Similarly, individuals experiencing a prolonged state of resource depletion can become exhausted and disengaged, and are thus less likely to restore their personal resources. Such a state can then lead to one experiencing low self-efficacy. As such, one who is low in resources may concurrently experience exhaustion, disengagement, and low self-efficacy and the cycle of experiencing these symptoms can perpetuate in the future.

There is an emerging body of evidence that challenges the assumption that TSE is an antecedent construct. For example, Holzberger, Philipp, and Kunter (2013) (2013) found using cross-lagged structural equation analyses that TSE did not longitudinally predict the two dimensions of student-reported teaching quality (cognitive activation and learning support). Rather, high student-reported teaching quality dimensions longitudinally predicted high TSE. Similarly, Praetorius and colleagues (2017) found using a cross-lagged autoregressive model that TSE did not longitudinally predict student-reported teaching quality (classroom management, cognitive activation, and learning support). Rather, high teaching quality longitudinally predicted high TSE. These two studies indicate that TSE can be considered an outcome construct. However, these empirical studies were very similar to each other in that they both studied German secondary mathematics teacher populations and used a very similar outcome measure. Thus, whether this finding would generalize to teachers with other characteristics (e.g., country, instructing educational level, and subject area) and to other outcome variables (e.g., dimensions of burnout) is yet unknown.

Some preliminary evidence indicates that burnout appears concurrently with and even precedes TSE. Brouwers and Tomic (2000) measured three dimensions of burnout (emotional exhaustion, personal accomplishment, and depersonalization) using the Maslach Burnout Inventory (MBI) over two time points at intervals of five months. They compared the model fit statistics within synchronous (construct from the same time point) and longitudinal models (construct from Time 1 predicting another construct at Time 2) and found that emotional exhaustion was associated with TSE at the same time point. Furthermore, TSE was associated with personal accomplishment at the same time point and predicted future levels of depersonalization. Brouwers, Evers, and Tomic (2001) reported that TSE predicted the three burnout dimensions in a closing sequence; namely, TSE predicted emotional exhaustion, which in turn predicted depersonalization, which in turn predicted personal accomplishment, which in turn predicted TSE. Together, these studies indicate that burnout can be both predicted by and predict TSE. However, their interpretations should be approached with caution as they drew causal conclusions though they used cross-sectional data. As such, in order to accurately determine the nature of the association between TSE and burnout, it is important to use a longitudinal study implementing a full-panel design that can test the reciprocal associations.

Autoregressive cross-lagged panel models (Campbell, 1963; Kenny, 1973; Kenny & Harackiewicz, 1979) using data from multiple time points allow us to test such reciprocal associations between the two constructs. These models have been previously used to study the directional nature of constructs, including job demands, job resources, burnout, and work engagement (Hakanen et al., 2008) and parental involvement and student mathematical achievement (Hong, Yoo, You, & Wu, 2010). This type of analytical approach has not been used so far to examine the directional association between self-efficacy and burnout. Thus, we use cross-lagged panel models based on three-wave longitudinal data to examine whether low TSE is concurrently associated with exhaustion and disengagement, whether low TSE causes exhaustion and disengagement, and/or whether exhaustion and disengagement causes low TSE.

Moderators and Covariates of Teacher Self-Efficacy and Burnout

The association between TSE and burnout has often been examined without considering factors that may moderate this relationship or impact the levels of the two constructs. Although studies have examined the stability of each of the constructs alone (e.g., Hakanen et al., 2008; Holzberger et al., 2013; Pas et al., 2012; Schaufeli et al., 2011), studies have not yet examined whether the TSE–burnout dimension associations are invariant across time and teacher demographic variables. Examining moderators such as time, gender, career stage, and instructing educational level are particularly important to assess the generalizability of the findings.

Additionally, there are patches of evidence suggesting that gender, years of teaching, and instructing education level may be covariates of TSE and the dimensions of burnout. More specifically, female teachers tend to have lower levels of TSE (Skaalvik & Skaalvik, 2007) and higher levels of burnout (Skaalvik & Skaalvik, 2007; Skaalvik & Skaalvik, 2017) than male teachers. Years of teaching experience is negatively associated with TSE (Skaalvik

& Skaalvik, 2007) and positively associated with burnout (Skaalvik & Skaalvik, 2009). TSE levels also seem to vary across instructing educational levels. Klassen and Chiu (2010) examined three domains of TSE (instructional strategies, classroom management, and student engagement) and found that kindergarten teachers have higher levels of classroom management self-efficacy than Grade 1 or 2 teachers and kindergarten teachers have higher levels of student engagement self-efficacy than teachers instructing higher grades. Evidence on whether levels of the dimensions of burnout differ across educational levels seems to be lacking.

Overview of the Current Study

Using an autoregressive cross-lagged panel model with longitudinal data on TSE and the two burnout dimensions (i.e., exhaustion and disengagement) collected at three time points, we aim to clarify the nature and the directionality of the associations between TSE and the burnout dimensions. Given the lack of evidence on the differences in the associations between TSE and the dimensions of burnout, we outline our hypotheses such that we do not expect different results between exhaustion and disengagement.

First, we assess the causal ordering of the TSE–burnout dimension associations by determining the nature of their associations across time and within a single time point. Specifically, we examine whether TSE and burnout are concurrently associated with one another, whether TSE predicts future burnout, and/or whether burnout predicts future TSE. We hypothesize TSE and burnout will be associated with each other at the same time point (H1). Moreover, we hypothesize that current TSE levels will predict future TSE levels (H2) and current burnout levels will predict future burnout levels (H3). Furthermore, we hypothesize current TSE levels will predict future burnout levels (H4) and current burnout levels will predict future TSE levels (H5).

Second, we examine the invariance of the structural paths in the final cross-lagged path model across four factors. Specifically, we examine whether the TSE–burnout dimension associations vary across (a) time, (b) gender, (c) career stages (i.e., early-, mid-, vs. late-career), and (d) instructing educational levels (i.e., elementary, middle, vs. secondary school). Given the limited number of studies exploring these questions, and thus without evidence that these associations may be moderated by these four factors, we tentatively hypothesize that we will reject the hypothesis that the strength of the TSE–burnout dimension associations will vary across time (H6), gender (H7a), career stages (H7b), and instructing educational levels (H7c).

Lastly, we examine the structural paths between TSE and the two dimensions of burnout, after controlling for the three covariates (i.e., gender, years of teaching experience, and instructing educational level). We hypothesize that the associations specified (i.e., H1-H5) will remain the same, even after controlling for the covariates (H8-H12).

Method

Participants and Procedure

A full panel design based on three time points was employed. At the first wave of data collection (Autumn, 2015), a convenient sample of 3002 Croatian teachers (82% female) across 135 state schools located in various parts of Croatia voluntarily participated in the study. At the time of initial data collection, they were, on average, 41.75 years old (SD = 10.44) and had, on average, 15.28 (SD = 10.50) years of teaching experience. Following Gu and Day's (2007) grouping procedure, teachers were split into three groups of experience levels in order to enable the test of invariance of the hypothesized associations across the career stages. Under this grouping, at Time 1 there were 802 early-career teachers (≤8 years of teaching experience), 1412 mid-career teachers (9-23 years of teaching experience), and 667 late-career teachers (≥24 years of teaching experience). Others did not provide information regarding years of teaching experience. In terms of instructing educational levels, 867 teachers taught at elementary level, 1056 at middle school level, 935 at secondary school level, and remaining teachers either did not respond to this item or taught students at multiple educational levels (e.g., both middle school and secondary school). Consistent with other studies on Croatian teachers (e.g., Jugović, Marušić, Pavin Ivanec, & Vizek Vidović, 2012), we did not ask for their ethnicity since Croatians are largely ethnically homogeneous.

Schools were recruited with the assistance of chiefs of the County Councils of School Psychologists (*n* = 12), who contacted the school psychologists under their supervision and informed them about the research project. After receiving consent from the school psychologists, the chiefs delivered to the research team the list of schools whose teachers agreed to voluntarily participate in the research. Approximately 50% of the teachers from the listed schools completed the questionnaire at the first assessment point, which is considerably higher than in previous studies on teachers (e.g., Mertler, 2003). For each of the three time points, questionnaires were sent to schools via postal service and distributed to the teachers by the school psychologists. After approximately two weeks, the school psychologists returned the completed questionnaires to the research team. Teacher responses over the three time points were matched using specially created codes known only to the teachers in order to preserve their anonymity.

Attrition Analysis

Of the initial sample (*N*=3002), 1525 (51%) teachers left at the second assessment point (Spring, 2016) and 1081 (36%) teachers left at the third assessment point (Autumn, 2016). Therefore, an attrition analysis was conducted to test the extent to which the teacher dropped out was related to either the covariates (i.e., gender, career stage, and educational level) or to the substantive variables (i.e., TSE, exhaustion, and disengagement). To test the sample structure in terms of gender and educational level across the three time points, a series of chi-squared tests was conducted. Using Time 1 as a baseline measure, male teachers were less likely than female teachers to participate at Time 2, $\chi^2(1) = 11.36$, $p < .01$, and at Time 3, $\chi^2(1) = 11.89$, $p < .01$. The ratio of male and female teachers did not change from Time 2 to Time 3, $\chi^2(1) = 1.00$, $p > .05$. Concerning the educational level, there was a higher number of elementary teachers and a smaller number of high school teachers at Time 3 than at Time 1, $\chi^2(2) = 40.49$, $p < .01$, and at Time 2, $\chi^2(2) = 28.13$. However, there was no difference in the ratio of teachers at the different educational levels between Time 1 and Time 2, $\chi^2\ (2) = 5.81$, $p > .05$.

Next, we tested whether teachers who dropped out at different time points differed in their years of teaching experience and the substantive variables (i.e., TSE, exhaustion, and disengagement). Series of *t*-tests were conducted to compare the teachers who participated or dropped out after different time points. Compared to teachers who participated at both Time 1 and 2, those who dropped out after Time 1 had slightly higher levels of TSE, $t(2944) = -2.09$, $p = .037$, $d = .08$, and lower levels of exhaustion, $t(2906) = 3.607$, $p = .001$, $d = .14$. There were no statistical differences between teachers who participated at all three time points with those who dropped out after Time 1 or Time 2. Lastly, a comparison of teachers who participated at all three time points with teachers who dropped out after Time 2 and teachers who dropped out after Time 1, again showed significant differences in their levels of TSE, $F(2, 2943) = 3.10$, $p = .045$, and exhaustion, $F(2, 2909) = 6.62$, $p = .001$. LSD post hoc analysis showed that teachers who participated at all three time points had higher initial levels of TSE ($p < .05$, $d = .10$) and lower initial levels of exhaustion ($p <.01$, $d = .14$) than teachers who dropped out after Time 1. In addition, teachers who dropped out after Time 2 also had higher initial levels of exhaustion when compared to teachers who dropped out after Time 1 ($p < .01$, $d = .16$). No other differences regarding substantive variables were found. Detailed information on teacher demographics and substantive variables across measurement occasions are presented in the Appendix.

The results of the attrition analysis indicated the justifiability of including the demographic covariates in the main analysis to control for their potential effects on the substantive variables. In addition, even though there were some statistically significant differences in TSE and exhaustion between completers and non-completers (which can be, at least in part, attributable to a large sample size and consequently great statistical power of this research), the effect sizes were quite small ($d < .20$; Cohen, 1988) and, thus, unlikely to seriously bias the results. Therefore, it was decided to proceed with the full information maximum likelihood procedure (FIML; Enders, 2010) in order to handle the missing data, which is an appropriate method to manage missing data in longitudinal studies (Jeličič, Phelps, & Lerner, 2009).

Measures

This study was part of a larger research project aimed to investigate teachers' emotion and emotion regulation, its personal and contextual antecedents, and effects on teacher functioning (Burić, 2019; Burić & Macuka, 2017; Burić, Penezić, & Sorić, 2017; Burić, Slišković, & Macuka, 2017; Burić, Slišković, & Penezić, 2019a; Burić, Slišković, & Penezić, 2019b; Slišković, Burić, & Macuka, 2016). To answer the research questions from this study, data on teacher demographics, TSE, and the burnout dimensions was used. The descriptive statistics and Cronbach alphas for these three groups of measures administered across the three time points are presented in Table 1.

INSERT TABLE 1 ABOUT HERE

TSE was measured using the *Teacher Self-efficacy Scale* (TSE; Schwarzer et al., 1999) which consists of 10 items assessing teachers' perception of their efficacy in job accomplishment, skill development, social interactions with students, and coping with job stress. Teachers gave their responses on a 4-point scale ranging from 1 (not at all true) to 4 (exactly true). An example item is: "I am convinced that I am able to successfully teach all relevant subject content to even the most difficult students."

To asses burnout, the *Oldenburg Burnout Inventory* (OLBI; Demerouti & Bakker, 2008) was administered. OLBI consists of two dimensions: *exhaustion* (n = 8; sample item: "During my work, I often feel emotionally drained") and *disengagement* (n = 8; sample item: "Lately, I tend to think less at work and do my job almost mechanically"). Teachers rated all items on a 4-point scale ranging from 1 (strongly disagree) to 4 (strongly agree).

Statistical Analyses

Five sets of statistical analyses were conducted using Mplus 8.0 (Muthén & Muthén, 1998–2017), consisting of two sets of preliminary analyses and three sets of main analyses. The preliminary analyses aimed to (a) establish the measurement invariance across time and (b) establish the measurement invariance across the three moderators (i.e., gender, career stage, and educational level). The main analysis aimed to (c) test the relevance of the first- and higher-order autoregressive and cross-lagged paths; (d) establish the structural invariance of the final cross-lagged path model across time and the three moderators; and (e) examine the stability of the structural paths of the final cross-lagged path model, after controlling for gender, years of experience, and educational levels.

The parameters in all models were estimated using the robust maximum-likelihood estimation method. The quality of model fit was assessed using four criteria: comparative fit index (CFI), Tucker-Lewis index (TLI), root-mean-square error of approximation (RMSEA), and standardized root-mean residual (SRMR). Traditionally, CFI and TLI values above .90

and .95 indicate acceptable and excellent fit, respectively (Hu & Bentler, 1999), while RMSEA values lower than .06 and SRMR values lower than .08 are indicative of good fit (Browne & Cudeck, 1993). To evaluate the measurement invariance, the Satorra-Bentler scaled chi-square difference test (TRd) was calculated to examine whether the difference was statistically non-significant, although a statistically significant value can also be attributed to its high sensitivity to large sample sizes (Marsh, Balla, & McDonald, 1998). Moreover, ΔCFI $\leq .01$ and ΔRMSEA $\leq .015$ criteria were used, with preference for models with lower values (Chen, 2007; Cheung & Rensvold, 2002). In addition, when choosing the best fitting structural model, AIC values were considered — an increase of AIC > 10 suggests a worse fitting and essentially an unacceptable model (Burnham & Anderson, 2002).

We examined the two dimensions of burnout separately in our analyses in order to model the multidimensional nature of the construct, to avoid problems with multicollinearity, and to reduce model complexity. Furthermore, even though the data used in this study have a hierarchical structure (i.e., teachers are nested within schools), all analyses were conducted at the teacher level only due to the negligible ICC1 values, which ranged from 0.003 to 0.012 for all substantive variables across all time points.

Preliminary analyses

Pearson correlations. Pearson correlation coefficients were calculated between TSE, exhaustion, and disengagement at all three time points.

Measurement invariance models. A necessary condition to be able to conduct the following main analyses is to establish the measurement invariance of the TSE–burnout dimension associations across time (i.e., three time points). The scale items were used as indicators of each of the three latent variables (i.e., TSE, exhaustion, and disengagement). The residuals of these scale items across the three time points were allowed to correlate with each other to control for systematic measurement errors (Marsh & Hau, 1996). As a further necessary condition for the main analyses, we also examined the measurement invariance of the models across the three moderators (i.e., gender, career stage, and educational level).

To establish a sufficient amount of measurement invariance of the latent constructs, the configural invariance and metric invariance must be achieved prior to testing the invariance of the structural paths (Byrne, 2012). A configural invariance model is less restrictive than a metric invariance model, as only invariance of the configuration of the associations between the latent constructs and their indicators has to be established. In a metric invariance model, the factor loadings are equivalent across tested moderators in addition to configural invariance (Byrne, 2012).

Main analyses.

Higher-order autoregressive and cross-lagged path models. In order to determine which type of model best describes the relationships between TSE and the two burnout dimensions, and so be able to examine H1-H5, two sets of four structural models were specified, tested, and compared to each other. The four structural models were: (a) a full-forward model, which includes both first- and higher-order stability and cross-lagged paths (M1); (b) a model, which includes first- and higher-order stability paths but only first-order cross-lagged paths (M2); (c) a model, which includes first-order stability paths and first- and higher-order cross-lagged paths (M3); and (d) a model, which includes only first-order stability and cross-lagged paths (M4). It should be noted that in each of the models, TSE was specified to correlate with a respective burnout dimension within a single time point. These models are depicted in Figure 1.

INSERT FIGURE 1 ABOUT HERE

Structural invariance models. To examine H6 and H7a-H7c, the invariance of the autoregressive and cross-lagged paths of the best fitting structural model (among M1-M4) were tested across time and the three moderators (i.e., gender, years of experience, and instructing educational levels). The models, in which stability and cross-lagged paths were set to be equal in size across these four factors, were compared to the baseline models where these structural paths were allowed to freely vary across these four factors.

Final cross-lagged path model with covariates. To examine H8 to H12, the three demographic variables were introduced as covariates in the best fitting structural models. More precisely, TSE and dimensions of burnout at Time 1 were regressed on teacher gender, instructing educational level, and years of teaching experience.

Results

Preliminary Analyses

Pearson correlations. Pearson correlation coefficients were calculated for the manifest variables assessed at all three time points. As can be seen in Table 1, TSE was negatively associated with both exhaustion and disengagement. This pattern of association remains stable both within a single time point and across time. Additionally, slightly higher levels of exhaustion were reported by female teachers at all three time points ($r = .11$, $r = .06$, and $r = .06$, respectively) and by more experienced teachers at Time 1 and Time 2 ($r = .07$ and $r = .07$, respectively). Lastly, more experienced teachers reported somewhat higher levels of TSE at Time 1 ($r = .07$).

Measurement invariance models. In order to test the measurement invariance of TSE–Exhaustion and TSE–Disengagement models across time and the three moderators, a series of models were tested. The fit statistics of these models are presented in Table 2 and Table 3. It should be noted that all the models, regardless of the imposed restrictions, demonstrated either excellent (i.e., RMSEA and SRMR) or acceptable fit to the data (i.e., CFI and TLI). It should be noted that traditional criteria are found to be overly strict for complex data (Heene, Hilbert, Draxler, Ziegler, & Bühner, 2011), as they were in this study.

INSERT TABLE 2 ABOUT HERE

The tests of TSE–Exhaustion models showed that, based on most criteria, the metric invariance models did not demonstrate worse fit than the configural invariance models, and these results held across time (TRd=74.21, Δ*df*=32, *p*<.01; ΔCFI = .001, ΔRMSEA=.000), gender (TRd=56.55, Δ*df*=54, *p*>.05; ΔCFI = .000, ΔRMSEA=.000), career stages (TRd=120.60, Δ*df*=108, *p*>.05; ΔCFI=.000 and ΔRMSEA=.000), and educational levels (TRd=125.77, Δ*df*=108, *p*>.05; ΔCFI=.000 and ΔRMSEA=.000). In addition, setting factor loadings to be equal in TSE–Disengagement models also did not result in any substantial loss in model fit when compared to the less restrictive configural models with regard to time (TRd=50.06, Δ*df*=32, *p*<.05; ΔCFI=.001 and ΔRMSEA=.000), gender (TRd=69.12, Δdf=54, *p*>.05; ΔCFI = .001, ΔRMSEA=.000), career stage (TRd=117.19, Δdf=108, *p*>.05; ΔCFI=.001 and ΔRMSEA=.000), or educational level (TRd=155.00, Δdf=108, *p*<.05; ΔCFI=.002 and ΔRMSEA=.001). Even though in some model comparisons statistically significant values of Sattora-Bentler scaled chi-square difference tests were obtained, they can be attributed to high sensitivity of the chi-square test to large sample sizes (Marsh et al., 1998) as it was the case in this study. Thus, it can be concluded that sufficient amount of measurement invariance was achieved across all analyzed moderators, which justified the subsequent tests of structural invariance.

INSERT TABLE 3 ABOUT HERE

Main Analyses

Higher-order autoregressive and cross-lagged path models. We tested the hypothesized structural associations depicted in Figure 1 via cross-lagged Structural Equation Modeling. As can be seen from Table 2, all tested TSE–Exhaustion models demonstrated similar model fit; however, fine-grained differences can be observed in AIC values. More specifically, in comparison with the full-forward model (M1), removing the second order cross-lagged paths (M2) did not result in a loss of model fit (ΔCFI = .000, ΔRMSEA = .000, ΔAIC = .214). When compared to the full-forward model (M1), both the model with the first- and higher-order cross-lagged paths but only the first-order stability paths (M3), and a model that includes only the first-order stability and cross-lagged paths (M4) had a worse fit, at least based on ΔAIC values (ΔCFI = .003, ΔRMSEA = .000, ΔAIC = 104.935 and ΔCFI = .003, ΔRMSEA = .000, ΔAIC =109.335, respectively). In addition, models M3 and M4 also had a worse fit when compared to the more parsimonious M2 model (ΔCFI = .003, ΔRMSEA = .000, ΔAIC = 104.721 and ΔCFI = .003, ΔRMSEA = .000, ΔAIC = 109.121, respectively). Considering a negligible difference between the full-forward model (M1) and the more parsimonious M2 model that includes both the first- and higher-order stability paths but only the first-order cross-lagged paths, the latter model was chosen as the best fitting one.

Similar conclusions could be drawn in regard to findings from TSE–Disengagement structural models. As results in Table 3 indicate, M2 model did not fit the data worse compared to the full-forward model (M1; ΔCFI = .000, ΔRMSEA = .000, ΔAIC = 4.825). However, a loss of the model fit was observed when contrasting M1 to more parsimonious M3 and M4 models (ΔCFI = .003, ΔRMSEA = .000, ΔAIC = 77.551 and ΔCFI = .003, ΔRMSEA = .000, ΔAIC = 102.201, respectively). Moreover, M3 and M4 models showed worse model fit when compared to M2 as well (ΔCFI = .003, ΔRMSEA = .000, ΔAIC = 72.726 and ΔCFI = .003, ΔRMSEA = .000, ΔAIC =97.376, respectively). As with the TSE–

Exhaustion models, the model that includes both the first- and higher-order stability paths but only the first-order cross-lagged path (M2) was chosen as the best fitting model due to its greater parsimony when compared to the full-forward model (M1).

In summary, the best fitting model was that which was in full support for H1 to H3 but in partial support for H4 and H5. Specifically, TSE and the burnout dimensions were indeed associated with each other at the same time point (H1), current TSE levels predicted future TSE levels (H2), and current burnout levels (both exhaustion and disengagement) predicted future burnout levels (H3). Current TSE levels predicted future burnout levels but only for one burnout dimension at one time interval (i.e., TSE_{T1} to $\text{Disengagement}_{\text{T2}}$), which partially supports H4. Furthermore, current burnout levels (both exhaustion and disengagement) predicted future TSE levels only for adjacent time points (i.e., Time 1 to Time 2 and Time 2 to Time 3), which partially supports H5.

Structural invariance models. After establishing the best fitting models representing the structural relationships between TSE and the two dimensions of burnout (i.e., M2), we tested whether the stability and cross-lagged paths were sufficiently invariant across the four factors (i.e., time, gender, career stage, and educational level). As expected, and as shown in Table 2 (Models 13 to 20), the model fit indices in the TSE–Exhaustion M2 model did not worsen when the paths were constrained across time (ΔCFI = .001 and ΔRMSEA = .000; H6), gender (ΔCFI = .000 and ΔRMSEA = .000; H7a), career stages (ΔCFI = .001 and ΔRMSEA = .001; H7b), and educational levels (ΔCFI = .000 and ΔRMSEA = .000; H7c). Similarly, as expected and as shown in Table 3 (Models 13 to 20), the model fit indices in the TSE–Disengagement M2 model also did not worsen when the paths were constrained across time (ΔCFI = .003 and ΔRMSEA = .001; H6), gender (ΔCFI = .000 and ΔRMSEA = .000; H7a), career stages (ΔCFI = .000 and ΔRMSEA = .001; H7b), and educational levels (ΔCFI = .000 and ΔRMSEA = .000; H7c). Thus, the longitudinal structural paths between TSE and the burnout dimensions seemed to be equivalent across time, gender, career stages, and instructing educational levels, which were in line with our expectations (H6 and H7a-H7c).

Final cross-lagged path model with covariates. The covariates of gender, years of teaching experience, and instructing educational levels were introduced in the best fitting structural M2 models as exogenous variables at Time 1. The model fit indices after introducing covariates are shown in Table 2 (Model 21) for the TSE–Exhaustion model and Table 3 (Model 21) for the TSE–Disengagement model, while their regression coefficients are presented in Table 4.

INSERT TABLE 4 ABOUT HERE

TSE correlated negatively with both dimensions of burnout at each time point, in support for H8. Namely, TSE correlated negatively at each time point with exhaustion (r_{T1} = -.43, p < .001; r_{T2} = -.33, p < .001; and r_{T3} = -.38, p < .001) and with disengagement (r_{T1} = -.56, p < .001; r_{T2} = -.46, p < .001; and r_{T3} = -.46, p < .001). Furthermore, current TSE levels predicted future TSE levels (β = .299 to .626) and current burnout levels predicted future burnout levels (β = .205 to .745 for exhaustion; β = .233 to .738 for disengagement), in support for H9 and H10, respectively.

In regard to the direction of prediction from TSE to the dimensions of burnout, we found partial support for H11. Namely, as shown in Table 4, we found that the only statistically significant path was TSE predicting disengagement from Time 1 to Time 2 (β = -.075, p < .05). In regard to the direction of prediction from the dimensions of burnout to TSE, we found full support for H12 as all possible paths within M2 model were statistically significant. Namely, as also shown in Table 4, exhaustion at Time 1 negatively predicted TSE at Time 2 (β = -.070, p < .05) and exhaustion at Time 2 negatively predicted TSE at Time 3

(β = -.130, p < .001). In addition, disengagement at Time 1 negatively predicted TSE at Time 2 (β = -.120, p < .001) and disengagement at Time 2 negatively predicted TSE at Time 3 (β = -.180, p < .001).

Considering the effects of the covariates on TSE and the burnout dimensions at Time 1 within the final cross-lagged path model, female teachers reported higher levels of exhaustion (β = .311, p < .001) but not disengagement (β = -.031, p >. 05). Teachers with more years of teaching experience had somewhat higher levels of both exhaustion (β = .090, p < .001) and disengagement (β = -.068, p < .01). Regarding the instructing educational level, middle school teachers reported lower levels of TSE than teachers instructing at other educational levels (β = -.254, p < .001 and β = -.250, p < .001 for the two models, respectively). Similarly, high school teachers reported lower levels of TSE than teachers instructing at other educational levels (β = -.185, p < .001 and β = -.182, p < .001, respectively). In addition, both middle- and high school teachers had higher levels of disengagement when compared to other groups of teachers (β = .239, p < .001 and β = .239, p < .001, respectively).

Discussion

The current study examined the associations between TSE and the two dimensions of burnout based on a three-wave panel design. Our aim was to examine whether current TSE levels were concurrently associated with current burnout levels as well as whether current TSE levels predicted future burnout levels and/or whether current burnout levels predicted future TSE levels. After establishing measurement invariance, we determined the best model describing the associations between TSE and the burnout dimensions and tested whether the models were invariant across time, gender, career stages of the teacher, and instructing educational levels. Additionally, we examined the associations between TSE and the dimensions of burnout after controlling for the covariates (gender, years of teaching experience, and instructing educational level).

As expected and consistent with previous review findings (Aloe et al., 2014; Zee & Koomen, 2016), TSE and both dimensions of burnout were concurrently associated with each other at each of the three time points both in models without and with covariates. This finding supports studies, which modelled self-efficacy and burnout at the same theoretical and empirical time point, including Skaalvik and Skaalvik (2010) who considered self-efficacy and burnout as correlates and simultaneously examined the two constructs' predictors (i.e., perceived school context) and outcome (i.e., job satisfaction). Furthermore, current TSE levels predicted future TSE levels and current burnout levels predicted future burnout levels. These findings are in line with findings on the relative stability of these construct levels over time (e.g., Hakanen et al., 2008; Holzberger et al., 2013; Pas et al., 2012; Schaufeli et al., 2011), which indicate that teachers will experience similar levels of TSE and burnout levels if no changes are made.

Contrary to expectation, TSE levels did not consistently predict future burnout levels. The best fitting model was one containing paths from TSE to the burnout dimensions but only for adjacent times. In a full model including the covariates, TSE preceded disengagement only at one of the two possible intervals, indicating that TSE, may to some extent, be reciprocally associated with disengagement. However, given the specificity of the finding to only one burnout dimension and the inconsistency of this finding across the time intervals, it seems premature to conclude that TSE precedes burnout.

On the other hand, burnout dimensions consistently predicted future TSE levels. The best fitting model was one containing paths from burnout dimensions to TSE but only for adjacent times. However, in a full model including the covariates, all relevant paths possible within this model were statistically significant. That is, both dimensions of burnout preceded

TSE at both time intervals. Overall, we found that experiencing burnout (both exhaustion and disengagement) more strongly colors one's future TSE levels than TSE colors future burnout levels. It may suggest that the potency of a negative experience (e.g., burnout) is greater than an emotionally relatively neutral construct (e.g., efficacy about one's ability to carry out tasks in their job) in affecting one's future states and experiences.

Our findings on the temporal order of the two constructs challenge the assumption that TSE always predicts burnout (e.g., Schwarzer & Hallum, 2008) but are in line with theory and previous empirical findings. According to Hobfoll's (1989) Conservation of Resource theory, prolonged exposure to stressors (e.g., lack of collegiality) results in resource depletion, which can be manifested as burnout symptoms. Such a state can hinder one's ability to fill their resources and thus negatively influences one's level of confidence and self-efficacy. Similarly, Byrne (1998) and Huberman (1993) have claimed that experiences of burnout symptoms can have negative effects on teachers, including their motivation, belief, and ability to perform well in their job. Our finding is also in line with empirical findings from two groups of researchers (Holzberger et al., 2013; Praetorius et al., 2017), who found using a German secondary mathematics teacher sample that TSE was predicted by (and does not predict) student-reported teaching quality. Their findings, too, challenged previous assumptions that TSE is a predictor by reporting that it was rather found to be an outcome variable.

Consistent with our expectations, the associations between TSE and the burnout dimensions did not vary depending on time, gender, career stage of the teacher, nor instructing educational level. Although the trajectory of the TSE and burnout levels throughout time and career stages have been noted previously (e.g., Holzberger et al., 2013; Klassen & Chiu, 2010; Praetorius et al., 2017), the steadfastness of the TSE–burnout associations across these four factors are new findings. That is, it seems that unless active strategies are undertaken, the effects of low TSE, burnout, and their consequences may carry over time whatever the gender, level of teaching experience a teacher has, and the educational level they instruct in.

The persistence of TSE–burnout associations over time highlights that the factors associated with burnout need to be addressed, especially when teachers are showing or beginning to show symptoms of burnout. The need to address these factors are further strengthened by Shoji and colleagues' (2016) meta-analytic finding that the strength of the association between self-efficacy and burnout are stronger among teachers than among other occupational groups. Some studies have recommended that strategies should be employed to increase TSE levels as they are a protective resource factor against negative outcomes (e.g., Schwarzer & Hallum, 2008). However, our findings challenge the logic of this strategy since it is the burnout dimensions that consistently predict TSE and not necessarily the other way around. Thus, deploying strategies to increase TSE levels may have minimal effects. Our challenge to the logic of this strategy is bolstered by the findings from Zee and Kooman (2016), who reported in their review that TSE does not directly predict teacher attrition. Rather, teachers with low self-efficacy experience burnout, which in turn leads them to quit their jobs. In this light, to tackle teacher attrition and other negative outcomes, strategies may need to focus on preventing and/or ameliorating the burnout symptoms rather than raising TSE levels. Otherwise, it is possible that unless external measures are brought in to fill their resources (e.g., changes in the school culture), teachers will continue to experience the burnout symptoms and also show other signs of resource depletion, including low self-efficacy, concurrently and in the future.

Theoretical and Practical Implications

Understanding the theoretical nature of a construct precedes any examination of its association with other constructs, as misrepresentation of the construct's nature can affect the conclusions one can draw from the findings. We observe that this nuanced understanding and representation are important in studying the construct of burnout. Although burnout is a multidimensional construct (Demerouti & Bakker, 2008; Maslach et al., 2001), some have examined the construct as a unidimensional one (e.g., Yu et al., 2015) to reduce model complexity and some have examined only one of its dimensions (e.g., Pas et al, 2012) perhaps to reduce survey administration time. When we examined TSE's associations with exhaustion and disengagement separately, we found that the nature of the associations differed between the two dimensions. As such, we recommend future studies to model burnout as a multidimensional construct and to draw conclusions about the construct in general, only if there is a consistency between the dimensions of the construct. Statistically recognizing a construct's multidimensionality will not only capture the nuanced detail in the findings, it will also subsequently help theory development.

Practically, teacher exhaustion and disengagement can be prevented and reduced with strategies that can be implemented at multiple levels, including at the individual, school, and teacher education program level. Teachers often find regulating their emotional resources difficult (Chang, 2009). Accordingly, pre-service and in-service teachers can learn to use effective cognitive emotion regulation strategies when experiencing negative events, such as positive reappraisal and putting events into perspective, to manage negative experiences (Burić, Penezić, & Sorić, 2017; Garnefski, Kraaij, & Spinhoven, 2001). Schools can also assist in the implementation of a variety of strategies. For example, Pas, Bradshaw, and Hershfeldt (2012) reported that burnout was negatively predicted by teachers' perceptions of preparedness in doing their job, teachers' perceptions of the collegial leadership, affiliation with the school and staff, and parent and student involvement. Accordingly, schools may find implementing strategies to address these four factors (two of which are also predictors of

TSE) helpful in preventing their teachers from burning out, which will not only benefit the teacher emotionally, but also the school financially and the students academically.

Teacher education programs can also integrate strategies and training as part of the curriculum. One particular strategy is to increase professional knowledge of the pre-service teachers. A study found that strengthening two aspects of professional knowledge (i.e., knowledge of learning and development, and knowledge of assessment) positively predicted a decrease in emotional exhaustion over time (Dicke, Parker, et al., 2015). A program focused on increasing classroom management skills may also be helpful as a study found that such training resulted in higher reports of well-being, including a reduction of emotional exhaustion (Dicke, Elling, Schmeck, & Leutner, 2015). These types of strategies may be helpful to prevent exhaustion and disengagement and enhance emotional well-being, which could form a part of pre-service teacher education and/or in-service professional development programs.

Limitations and Future Directions

Limitations exist within the study that future studies may wish to examine, such as testing the generalizability of our findings. The demographical characteristics of the current teacher sample are similar to that of the national teacher population in Croatia (Organization for Economic Cooperation and Development, 2015). Furthermore, the associations between self-efficacy and burnout may not vary across cultures (Shoji et al., 2016). Nevertheless, future studies may still wish to replicate our findings in different cultures as our sample was from Croatia. The Croatian educational system can be seen as unique in that it is characterized by humanistic values and didactic orientation, which has been undergoing transition and change within the globalization and European integration process (Cain & Milovic, 2010). Furthermore, the teachers in our study reported higher levels of self-efficacy when compared to teachers of other countries from previous studies (e.g., Holzberger et al., 2013; Praetorius et al., 2017; Schwarzer & Hallum, 2008). This result may be due to the truly higher levels of self-efficacy of Croatian teachers or due to these teachers using particularly high levels of self-serving strategies while answering the scale items (e.g., engaging in high levels of socially desirable responding). Whatever the case, replicating the study using other country samples may be beneficial.

Furthermore, we collected self-report data for TSE and the burnout dimensions given that it is the most direct and common way to capture one's inner states or subjective beliefs. Some may argue that self-report data can be subject to social desirability and bias (Paulhus, 2002) and may not be directly translated into actual behavior. In this light, future studies may benefit from collecting other sources of data, including other-reported (e.g., principal-report of teachers' self-efficacy) and behavioral data (classroom observation measures of belief practices), as well as data from teachers from other countries.

We collected our data using a field study design, whereby teachers reported their levels of the burnout dimensions and TSE as they have been experiencing them at the three time points at six-month intervals. Greater time intervals and greater number of assessment points may clarify the strength of the temporal effect of our findings. Future studies may also seek to strengthen evidence on the directional relationship of the two constructs by using an intervention study design, whereby strategies are implemented to reduce teacher levels of exhaustion and disengagement so to examine whether the strategy has a longitudinal predictive effect on TSE levels.

Lastly, the attrition analysis showed that teachers with somewhat higher levels of self-efficacy and lower levels of exhaustion tended to drop out from this study at earlier time points of the data collection, implying the possibility that our results may have been biased. That is, a greater tendency of the better-adjusted teachers to drop-out from the study at earlier points may have resulted in a restriction of range of the data and consequently the reduction in size of the established structural associations. Future studies may wish to employ additional strategies, such as offering increased amounts of incentives for each additional time point and ensuring that data is collected outside of marking and reporting periods, to ensure a relatively high retention rate.

Conclusion

In sum, this study clarified our theoretical and empirical understanding of the causal order and nature of the associations between self-efficacy and the dimensions of burnout (exhaustion and disengagement) in teachers. Specifically, our study findings contribute to previous study findings and arguments (Holzberger et al., 2013; Praetorius et al., 2017) that TSE should not necessarily be examined as an antecedent variable but as a consequential variable. Understanding the causal order of teacher-relevant constructs will help us not only in developing richer theory but also correctly implementing strategies to assist our teachers in their professional lives.

References

Aloe, A. M., Amo, L. C., & Shanahan, M. E. (2014). Classroom management self-efficacy and burnout: A multivariate meta-analysis. *Educational Psychology Review*, *26*, 101–126.

Arens, A. K., & Morin, A. J. S. (2016). Relations between teachers' emotional exhaustion and students' educational outcomes. *Journal of Educational Psychology*, *108*, 800–813.

Bakker, A. B., & Schaufeli, W. B. (2000). Burnout contagion processes among teachers. *Journal of Applied Social Psychology*, *30*, 2289–2308.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, *84*, 191–215.

Betoret, F. D. (2009). Self-efficacy, school resources, job stressors and burnout among Spanish primary and secondary school teachers: a structural equation approach. *Educational Psychology Review*, *29*, 45–68.

Brouwers, A., Evers, W. J. G., & Tomic, W. (2001). Self-efficacy in eliciting social support and burnout among secondary-school teachers. *Journal of Applied Social Psychology*, *31*, 1474–1491.

Brouwers, A., & Tomic, W. (2000). A longitudinal study of teacher burnout and perceived self-efficacy in classroom management. *Teaching and Teacher Education*, *16*, 239–253.

Browne, M.W., & Cudeck, R. (1992). Alternative ways of assessing model fit. *Sociological Methods & Research, 21*, 230–258.

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*, 112–129.

Burić, I. (2019). The role of emotional labor in explaining teachers' enthusiasm and students' outcomes: A multilevel mediational analysis. *Learning and Individual Differences*, *70*, 12–20.

Burić, I. & Macuka, I. (2017). Self-efficacy, emotions and work engagement among teachers: A two wave cross-lagged analysis. *Journal of Happiness Studies,* 1-17.

Burić, I., Penezić, Z., & Sorić, I. (2017). Regulating emotions in the teacher's workplace: Development and initial validation of the Teacher Emotion-Regulation Scale. *International Journal of Stress Management, 24*, 217-246.

Burić, I., Slišković, A., & Macuka, I. (2017). A mixed-method approach to the assessment of teachers' emotions: Development and validation of the Teacher Emotion Questionnaire. *Educational psychology*, 1-25.

Burić, I., Slišković, A., & Penezić, Z. (2019a). A two-wave panel study on teachers' emotions and emotional-labour strategies. *Stress and Health: Journal of the International Society for the Investigation of Stress*, *35*(1), 27–38.

Burić, I., Slišković, A., & Penezić, Z. (2019b). Understanding teacher well-being: a cross-lagged analysis of burnout, negative student-related emotions, psychopathological symptoms, and resilience. *Educational Psychology, 39*, 1136-1155.

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

Byrne, J. J. (1998). Teacher as hunger artist: Burnout: Its causes, effects, and remedies. *Contemporary Education, 69(*2), 86–91.

Byrne, M. B. (2012). *Structural Equation Modeling with Mplus.* NY: Routledge.

Cain, T., & Milovic, S. (2010). Action research as a tool of professional development of advisers and teachers in Croatia. *European Journal of Teacher Education*, *33*, 19–30.

Campbell, D. T. (1963). From description to experimentation: Interpreting trends as quasi-experiments. In C. W. Harris (Ed.), *Problems in measuring change* (pp. 212–242). Madison, WI: University of Wisconsin Press.

Chang, M.-L. (2009). An appraisal perspective of teacher burnout: Examining the emotional work of teachers. *Educational Psychology Review*, *21*, 193–218.

Chen, F.F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling, 14*, 464–504.

Cherniss, C. (2017). Role of professional self-efficacy in the etiology and amelioration of burnout. In W. B. Schaufeli, C. Maslach, & T. Marek (Eds.), *Professional burnout: Recent developments in theory and research* (pp. 135–149). London, England: Routledge.

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

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*, 2nd Edition. Hillsdale: Lawrence Erlbaum.

Darling-Hammond, L. (2010). Recruiting and retaining teachers: Turning around the race to the bottom in high-need schools. *Journal of Curriculum and Instruction*, *4*, 16–32.

Demerouti, E., & Bakker, A.B. (2008). The Oldenburg Burnout Inventory: A good alternative to measure burnout and engagement. In J. Halbesleben (Ed.), *Stress and burnout in health care* (pp. 65-78). NY: Nova Sciences Hauppauge.

Demerouti, E., Bakker, A. B., Vardakou, I., & Kantas, A. (2003). The convergent validity of two burnout instruments. *European Journal of Psychological Assessment*, *19*, 12–23.

Dicke, T., Elling, J., Schmeck, A., & Leutner, D. (2015). Reducing reality shock: The effects of classroom management skills training on beginning teachers. *Teaching and Teacher Education*, *48*, 1–12.

Dicke, T., Parker, P. D., Holzberger, D., Kunina-Habenicht, O., Kunter, M., & Leutner, D. (2015). Beginning teachers' efficacy and emotional exhaustion: Latent changes, reciprocity, and the influence of professional knowledge. *Contemporary Educational Psychology*, *41*, 62–72.

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?* (No. Policy Paper 19). UNESCO. Retrieved from http://unesdoc.unesco.org/images/0023/002327/232721e.pdf

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

Fackler, S., & Malmberg, L.-E. (2016). Teachers' self-efficacy in 14 OECD countries: Teacher, student group, school and leadership effects. *Teaching and Teacher Education*, *56*, 185–195.

Garnefski, N., Kraaij, V., & Spinhoven, P. (2001). Negative life events, cognitive emotion regulation and emotional problems. *Personality and Individual Differences*, *30*, 1311–1327.

Grayson, J. L., & Alvarez, H. K. (2008). School climate factors relating to teacher burnout: A mediator model. *Teaching and Teacher Education*, *24*, 1349–1363.

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

Hakanen, J. J., Schaufeli, W. B., & Ahola, K. (2008). The Job Demands-Resources model: A three-year cross-lagged study of burnout, depression, commitment, and work engagement. *Work & Stress*, *22*, 224–241.

Heene, M., Hilbert, S., Draxler, C., Ziegler, M., & Bühner, M. (2011). Masking misfit in confirmatory factor analysis by increasing unique variances: A cautionary note on the usefulness of cutoff values of fit indices. *Psychological Methods, 16*, 319-336.

Hobfoll, S. E. (1989). Conservation of resources. A new attempt at conceptualizing stress. *The American Psychologist*, *44*, 513–524.

Holzberger, D., Philipp, A., & Kunter, M. (2013). How teachers' self-efficacy is related to instructional quality: A longitudinal analysis. *Journal of Educational Psychology*, *105*, 774–786.

Hong, S., Yoo, S.-K., You, S., & Wu, C.-C. (2010). The reciprocal relationship between parental involvement and mathematics achievement: Autoregressive cross-lagged modeling. *Journal of Experimental Education*, *78*, 419–439.

Hu, L. & Bentler, P.M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1-55.

Huberman, M. (1993). Burnout in teaching careers. *European Education*, *25*(3), 47–69.

Ingersoll, R. M. (2002). The teacher shortage: A case of wrong diagnosis and wrong prescription. *NASSP Bulletin*, *86*(631), 16–31.

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

Jensen, B., Sandoval-Hernández, A., Knoll, S., & Gonzalez, E. J. (2012). *The experience of new teachers: Results from TALIS 2008*. OECD Publishing. http://dx.doi.org/10.1787/9789264120952-en

Johnson, S., Cooper, C., Cartwright, S., Donald, I., Taylor, P., & Millet, C. (2005). The experience of work-related stress across occupations. *Journal of Managerial Psychology*, *20*, 178–187.

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*, 1195-1199.

Jugović, I., Marušić, I., Pavin Ivanec, T., & Vizek Vidović, V. (2012). Motivation and personality of preservice teachers in Croatia. *Asia-Pacific Journal of Teacher Education*, *40*, 271–287.

Kenny, D. A. (1973). Cross-lagged and synchronous common factors in panel data. In A. S. Goldberger & O. D. Duncan (Eds.), *Structural equation models in the social sciences* (pp. 153–167). NY: Seminar Press.

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

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*, 741–756.

Kunter, M., Kleickmann, T., Klusmann, U., & Richter, D. (2013). The development of teachers' professional competence. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss, & M. Neubrand (Eds.), *Cognitive activation in the mathematics classroom and professional competence of teachers: results from the COACTIV project* (pp. 63–77). Boston, MA: Springer US.

Lamude, K. G., & Scudder, J. (1992). Resistance in the college classroom: Variations in students' perceived strategies for resistance and teachers' stressors as a function of students' ethnicity. *Perceptual and Motor Skills*, *75*, 615–626.

Leiter, M. P. (1992). Burnout as a crisis in self-efficacy: Conceptual and practical implications. *Work & Stress*, *6*, 107–115.

Leithwood, K. A., Menzies, T., Jantzi, D., & Leithwood, J. (1999). Teacher burnout: A critical challenge for leaders. In A. M. Huberman (Ed.), *Understanding and preventing teacher burnout: A sourcebook of international research and practice* (pp. 85–114). NY: Cambridge University Press.

Marsh, H. W., Balla, J. R., & McDonald, R. P. (1988). Goodness-of-fit indexes in confirmatory factor analysis: The effect of sample size. *Psychological Bulletin, 103,* 391– 410.

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

Martin, N. K., Sass, D. A., & Schmitt, T. A. (2012). Teacher efficacy in student engagement, instructional management, student stressors, and burnout: A theoretical model using in-class variables to predict teachers' intent-to-leave. *Teaching and Teacher Education*, *28*, 546–559.

Maslach, C., Schaufeli, W. B., & Leiter, M. P. (2001). Job burnout. *Annual Review of Psychology*, *52*, 397–422.

Mertler, C.A. (2003). Patterns of response and nonresponse from teachers to traditional and web surveys. *Practical Assessment, Research, & Evaluation, 8*(22), 1-9.

Muthén, L.K. & Muthén, B.O. (1998-2017). *Mplus User's Guide*. Eighth Edition. Los Angeles, CA: Muthén & Muthén.

Organization for Economic Cooperation and Development. (2005). *Teachers matter: Attracting, developing and retaining effective teachers*. Paris, France: OECD.

Organization for Economic Cooperation and Development (2015). *Country profile: Croatia*. Retrieved from http://www.oecd.org/education/school/TALIS-Country-profile-Croatia.pdf.

Pas, E. T., Bradshaw, C. P., & Hershfeldt, P. A. (2012). Teacher- and school-level predictors of teacher efficacy and burnout: Identifying potential areas for support. *Journal of School Psychology*, *50*, 129–145.

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). Mahwah, NJ: Erlbaum.

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*. NY: Routledge.

Schaufeli, W. B., Maassen, G. H., Bakker, A. B., & Sixma, H. J. (2011). Stability and change in burnout: A 10-year follow-up study among primary care physicians. *Journal of Occupational and Organizational Psychology, 84*, 248–267.

Schonfeld, I. S. (2001). Stress in 1st-year women teachers: The context of social support and coping. *Genetic, Social, and General Psychology Monographs*, *127*, 133–168.

Schwarzer, R., & Hallum, S. (2008). Perceived teacher self-efficacy as a predictor of job stress and burnout: Mediation analyses. *Applied Psychology*, *57*, 152–171.

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

Shoji, K., Cieslak, R., Smoktunowicz, E., Rogala, A., Benight, C. C., & Luszczynska, A. (2016). Associations between job burnout and self-efficacy: A meta-analysis. *Anxiety, Stress, and Coping*, *29*, 367–386.

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*, 611-625.

Skaalvik, E. M., & Skaalvik, S. (2009). Does school context matter? Relations with teacher burnout and job satisfaction. *Teaching and Teacher Education, 25*, 518–524.

Skaalvik, E. M., & Skaalvik, S. (2010). Teacher self-efficacy and teacher burnout: A study of relations. *Teaching and Teacher Education*, *26*, 1059–1069.

Skaalvik, E. M., & Skaalvik, S. (2017). Dimensions of teacher burnout: Relations with potential stressors at school. *Social Psychology of Education: An International Journal, 20*, 775–790.

Slišković, A., Burić, I., & Macuka, I. (2016). The voice of Croatian elementary school teachers: Qualitative analysis of the teachers' perspective on their profession. *Teachers and Teaching: Theory and Practice*, *23*, 518-531.

Sutcher, L., Darling-Hammond, L., & Carver-Thomas, D. (2016). *A coming crisis in teaching? Teacher supply, demand, and shortages in the US*. Palo Alto, CA: Learning Policy Institute.

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

Tschannen-Moran, M., & Hoy, A. W. (2007). The differential antecedents of self-efficacy beliefs of novice and experienced teachers. *Teaching and Teacher Education, 23*, 944–956.

UNESCO Institute for Statistics. (2016). *The world needs almost 69 million new teachers to reach the 2030 education goals* (No. 39). Retrieved from 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

Vandenberghe, R., & Huberman, A. M. (1999). *Understanding and preventing teacher burnout: A sourcebook of international research and practice*. Cambridge, England: Cambridge University Press.

Vlahović-Štetić, V., & Vizek Vidović, V. (2005). Obrazovanje učitelja i nastavnika u Hrvatskoj [Education of teachers in Croatia]. In V. Vizek Vidović (Ed.), *Cjeloživotno obrazovanje učitelja i nastavnika: višestruke perspektive [Lifelong learning of teachers: multiple perspectives]* (pp. 69–91). Zagreb: Institute for Social Research.

Voss, T., Wagner, W., Klusmann, U., Trautwein, U., & Kunter, M. (2017). Changes in beginning teachers' classroom management knowledge and emotional exhaustion during the induction phase. *Contemporary Educational Psychology*, *51*, 170–184.

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.

Yu, X., Wang, P., Zhai, X., Dai, H., & Yang, Q. (2015). The effect of work stress on job burnout among teachers: The mediating role of self-efficacy. *Social Indicators Research,* *122*, 701–708.

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*, 981–1015.

Table 1

Descriptive Statistics, Cronbach Alphas, and Correlations for Demographics, Self-Efficacy, Exhaustion, and Disengagement

	Variable	*M*	*SD*	Skewness (*SE*)	Kurtosis (*SE*)	2	3	4	5	6	7	8	9	10	11
1	Gender	n/a	n/a	n/a	n/a	.05*	.02	.02	-.01	.11**	.06*	.06*	.05	-.04	-.05
2	Experience (years)	15.28	10.50	0.50(0.45)	-0.68(0.09)	(n/a)	.03	.07**	.03	.07**	.07**	.00	.02	-.04	.05
3	Self-efficacy T1	3.37	0.40	-0.06(0.05)	0.79(0.09)		(.84)	.57**	.62**	-.38**	-.32**	-.34**	-.43**	-.38**	-.39**
4	Self-efficacy T2	3.33	0.41	-0.39(0.06)	0.43(0.13)			(.86)	.61**	-.31**	-.40**	-.33**	-.36**	-.49**	-.38**
5	Self-efficacy T3	3.29	0.44	-0.31(0.08)	0.14(0.15)				(.88)	-.35**	-.40**	-.48**	-.40**	-.44**	-.53**
6	Exhaustion T1	2.22	0.51	0.11(0.05)	0.16(0.09)					(.84)	.69**	.62**	.66**	.51**	.45**
7	Exhaustion T2	2.17	0.48	0.14(0.06)	0.34(0.13)						(.84)	.72**	.52**	.70**	.56**
8	Exhaustion T3	2.13	0.51	0.27(0.08)	0.58(0.15)							(.86)	.50**	.56**	.71**
9	Disengagement T1	2.01	0.47	0.21(0.05)	0.19(0.09)								(.76)	.69**	.63**
10	Disengagement T2	2.02	0.47	0.19(0.06)	0.31(0.18)									(.80)	.73**
11	Disengagement T3	2.01	0.48	0.23(0.08)	0.39(0.15)										(.80)

Note. Time = time point. SE = Standard Error. Cronbach αs are in parentheses. ** $p < .01$; * $p < .05$.

Table 2

Fit Statistics of TSE–Exhaustion Models

Model number	Model type	χ^2 (df)	CFI	TLI	RMSEA (90% C.I.)	SRMR	AIC
Measurement Invariance Models							
1	Configural invariance over time	3890.42 (1302)	.927	.920	.026 (.025, .027)	.055	163917.771
2	Metric invariance over time	3964.87 (1334)	.926	.920	.026 (.025, .027)	.055	163934.902
3	Configural invariance across gender	5648.33 (2604)	.917	.909	.028 (.027, .029)	.061	161887.901
4	Metric invariance across gender	5701.84 (2658)	.917	.911	.028 (.027, .029)	.063	161842.781
5	Configural invariance across career stages	6911.11 (3906)	.916	.908	.028 (.027, .029)	.063	157214.476
6	Metric invariance across career stages	7029.55 (4014)	.916	.910	.028 (.027, .029)	.068	157130.671
7	Configural invariance across educational levels	6905.32 (3906)	.915	.907	.028 (.027, .029)	.065	155966.735
8	Metric invariance across educational levels	7028.01 (4014)	.915	.909	.028 (.027, .029)	.071	155891.269
Higher-Order Autoregressive and Cross-Lagged Path Models							
9	M1 full-forward model; includes both first- and higher-order stability and cross-lagged paths	3964.87 (1334)	.926	.920	.026 (.025, .027)	.055	163934.902
10	M2 includes first- and higher order stability paths but only first-order cross-lagged paths	3968.84 (1336)	.926	.920	.026 (.025, .027)	.056	163935.116
11	M3 includes first-order stability paths and first- and higher-order cross-lagged paths	4063.11 (1336)	.923	.917	.026 (.025, .027)	.059	164039.837
12	M4 includes only first-order stability and cross-lagged paths	4069.53 (1338)	.923	.917	.026 (.025, .027)	.060	164044.237
Structural Invariance Models							
13	M2 with stability and cross-lagged paths unconstrained over time	3968.84 (1336)	.926	.920	.026 (.025, .027)	.056	163935.116
14	M2 with stability and cross-lagged paths constrained over time	3992.41 (1340)	.925	.920	.026 (.025, .027)	.057	163954.523
15	M2 with stability and cross-lagged paths unconstrained across gender	5779.71 (2664)	.915	.909	.028 (.027, .029)	.065	161918.355
16	M2 with stability and cross-lagged paths constrained across gender	5803.02 (2674)	.915	.909	.028 (.027, .029)	.068	161923.657
17	M2 with stability and cross-lagged paths unconstrained across career stages	7107.06 (4023)	.914	.908	.028 (.027, .029)	.070	157204.821
18	M2 with stability and cross-lagged paths constrained across career stages	7118.07 (4043)	.914	.909	.028 (.027, .029)	.071	157179.071
19	M2 with stability and cross-lagged paths unconstrained across educational levels	7100.01 (4023)	.913	.907	.028 (.027, .029)	.072	155959.750
20	M2 with stability and cross-lagged paths constrained across educational levels	7125.14 (4043)	.913	.907	.028 (.027, .029)	.074	155950.265
Final Cross-Lagged Path Model with Covariates							
21	M2 with effects of covariates at T1	4734.11 (1544)	.907	.901	.027 (.026, .028)	.055	153390.275

Table 3

Fit Statistics of TSE–Disengagement Models

Model number	Model type	χ^2 (df)	CFI	TLI	RMSEA (90% C.I.)	SRMR	AIC
Measurement Invariance Models							
1	Configural invariance over time	3258.91 (1302)	.939	.933	.022 (.021, .023)	.041	168792.747
2	Metric invariance over time	3310.29 (1334)	.938	.934	.022 (.021, .023)	.043	168781.784
3	Configural invariance across gender	4961.49 (2604)	.929	.922	.025 (.024, .026)	.049	166716.990
4	Metric invariance across gender	5028.78 (2658)	.929	.923	.025 (.024, .026)	.054	166686.784
5	Configural invariance across career stages	6288.81 (3906)	.927	.919	.025 (.024, .026)	.053	162087.564
6	Metric invariance across career stages	6404.85 (4014)	.926	.921	.025 (.024, .026)	.059	161999.484
7	Configural invariance across educational levels	6328.99 (3906)	.924	.916	.026 (.024, .027)	.055	160586.369
8	Metric invariance across educational levels	6483.19 (4014)	.922	.917	.025 (.024, .027)	.061	160542.994
Higher-Order Autoregressive and Cross-Lagged Path Models							
9	M1 full-forward model; includes both first- and higher-order stability and cross-lagged paths	3310.29 (1334)	.938	.934	.022 (.021, .023)	.043	168781.784
10	M2 includes first- and higher order stability paths but only first-order cross-lagged paths	3318.63 (1336)	.938	.934	.022 (.021, .023)	.043	168786.609
11	M3 includes first-order stability paths and first- and higher-order cross-lagged paths	3384.43 (1336)	.936	.931	.023 (.022, .024)	.045	168859.335
12	M4 includes only first-order stability and cross-lagged paths	3408.24 (1338)	.935	.931	.023 (.022, .024)	.047	168883.985
Structural Invariance Models							
13	M2 with stability and cross-lagged paths unconstrained over time	3318.63 (1336)	.938	.934	.022 (.021, .023)	.043	168786.609
14	M2 with stability and cross-lagged paths constrained over time	3424.45 (1340)	.935	.930	.023 (.022, .024)	.047	168896.246
15	M2 with stability and cross-lagged paths unconstrained across gender	5161.07 (2664)	.925	.920	.025 (.024, .026)	.054	166821.823
16	M2 with stability and cross-lagged paths constrained across gender	5187.19 (2674)	.925	.919	.025 (.024, .027)	.057	166830.096
17	M2 with stability and cross-lagged paths unconstrained across career stages	6547.65 (4023)	.922	.917	.026 (.024, .027)	.060	162141.544
18	M2 with stability and cross-lagged paths constrained across career stages	6565.36 (4043)	.922	.917	.025 (.024, .027)	.062	162124.569
19	M2 with stability and cross-lagged paths unconstrained across educational levels	6628.12 (4023)	.918	.913	.026 (.025, .027)	.064	160687.434
20	M2 with stability and cross-lagged paths constrained across educational levels	6648.13 (4043)	.918	.913	.026 (.025, .027)	.064	160675.299
Final Cross-Lagged Path Model with Covariates							
21	M2 with effects of covariates at T1	4137.231 (1544)	.917	.912	.024 (.024, .025)	.044	158077.659

Table 4

Standardized Factor Loadings and Path Coefficients of Final Cross-Lagged Path Models with Covariates

	TSE–Exhaustion model		TSE–Disengagement model	
Parameter estimate	TSE	Exhaustion	TSE	Disengagement
Factor loadings	.540-.703***	.461-.792***	.546-.709***	.337-.700***
Stability paths				
T1→T2	.626***	.745***	.582***	.738***
T2→T3	.335***	.623***	.299***	.627***
T1→T3	.416***	.205***	.397***	.233***
Cross-lagged effects	TSE → Exhaustion	Exhaustion → TSE	TSE → Disengagement	Disengagement → TSE
T1→T2	-.046	-.070*	-.075*	-.120***
T2→T3	-.044	-.130***	-.007	-.180***
Effects of covariates at T1				
Gender	.036	.311***	.034	-.031
Experience (years)	.028	.09***	.026	.068**
Middle school level	-.254***	.091	-.250***	.239***
High school level	-.185***	.071	-.182***	.239***

Note. Gender is coded 0=male, 1=female; Middle school level is coded 1=middle school teachers, 0=elementary and high school teachers; High school level is coded 1=high school teachers, 0=elementary and middle school teachers.

Effects of gender and educational levels were based on StdY standardization (i.e., standardization of dependent variables only) with Mplus. All other effects are based on StdYX standardization.

**p*<.05, ***p*<.01, ****p*<.001.

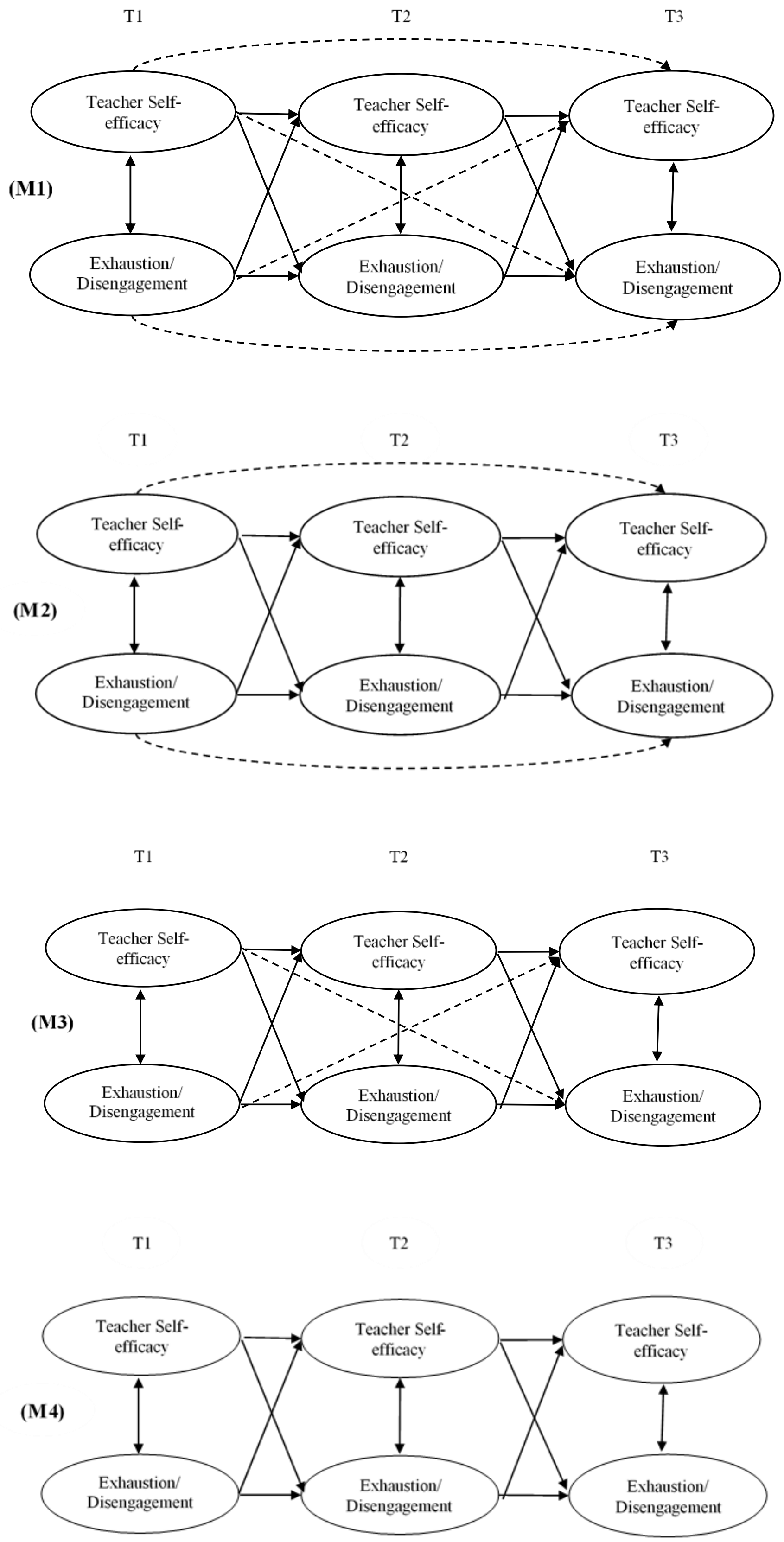

Figure 1. Hypothesized structural models.

Note. Higher-order paths are presented in dashed lines.

Appendix

Comparison Across Time and Comparison on Variables Between Completers and Noncompleters

Table A1

Teachers' Demographics Across Time

Variable		Time 1	Time 2	Time 3
		N (%)	*N* (%)	*N* (%)
Gender	Male	492 (16.6)	202 (13.4)	110 (12.3)
	Female	2474 (83.4)	1308 (86.6)	784 (87.7)
Educational level	Class teachers	867 (30.3)	473 (32.4)	325 (37.6)
	Middle-school teachers	1056 (36.9)	530 (36.3)	339 (39.2)
	High-school teachers	935 (32.7)	456 (31.3)	200 (23.1)
Career stage	Early Career	400 (27.2)	400 (27.2)	213 (24.3)
	Middle Career	1412 (48.9)	756 (51.3)	482 (55.0)
	Late Career	667 (23.1)	317 (21.5)	181 (20.7)

Note. Only valid responses are shown.

Table A2

Comparison on Variables Between Completers and Non-Completers

	Time 1			
Variable	Experience	TSE	Exhaustion	Disengagement
	M (SD)	*M (SD)*	*M (SD)*	*M (SD)*
Completers (T1+T2)	15.24 (10.17)	3.35 (0.39)	2.25 (0.51)	2.01 (0.46)
Non-completers (T1)	15.33 (10.82)	3.38 (0.41)	2.18 (0.51)	2.00 (0.48)
Completers (T1+T2+T3)	15.50 (9.89)	3.34 (0.40)	2.25 (0.50)	2.01 (0.45)
Non-completers (T1+T2)	14.86 (10.57)	3.37 (0.38)	2.26 (0.52)	2.02 (0.46)
	Time 2			
	Experience	TSE	Exhaustion	Disengagement
	M (SD)	*M (SD)*	*M (SD)*	*M (SD)*
Completers (T1+T2+T3)	16.10 (9.89)	3.32 (0.41)	2.18 (0.47)	2.03 (0.46)
Non-completers (T1+T2)	15.79 (10.75)	3.36 (0.41)	2.16 (0.50)	2.01 (0.49)

Note. TSE = Teacher Self-Efficacy; T= Time point.