Kernel Causality Among Teacher Self-Efficacy, Job Satisfaction, School Climate, and Workplace Well-Being and Stress in TALIS

Xue Zhang1*, Chunyang Zhao2, Yuqiao Xu3, Shanhua Liu1 and Zhihui Wu1

1 China Institute of Rural Education Development, Northeast Normal University, Changchun, China, 2 School of Mathematics and Statistics, Northeast Normal University, Changchun, China, 3 Faculty of Education, Northeast Normal University, Changchun, China

Teachers play an important role in the educational system. Teacher self-efficacy, job satisfaction, school climate, and workplace well-being and stress are four individual characteristics shown to be associated with tendency to turnover. In this article, data from the Teaching and Learning International Survey (TALIS) 2018 teacher questionnaire are analyzed, with the goal to understand the interplay amongst these four individual characteristics. The main purposes of this study are to (1) measure extreme response style for each scale using unidimensional nominal response models, and (2) investigate the kernel causal paths among teacher self-efficacy, job satisfaction, school climate, and workplace well-being and stress in the TALIS-PISA linked countries/economies.

Our findings support the existence of extreme response style, the rational non-normal distribution assumption of latent traits, and the feasibility of kernel causal inference in the educational sector. Results of the present study inform the development of future correlational research and policy making in education.

Keywords: kernel causality, generalized measure of correlation, nominal response model, extreme response style, TALIS

INTRODUCTION

In recent years, teacher turnover has become an increasingly prominent issue surrounding the topic of teaching quality. It was found that a high sense of teacher self-efficacy has an indirect effect on later job satisfaction via the mediational role of engagement (Granziera and Perera, 2019), controls professional stress (Bangs and Frost, 2012), increases teacher well-being (Smylie, 1988), and reduces quitting intentions (Tschannen-Moran and Hoy, 2001; Wang et al., 2015). Job satisfaction is considered as a critical predictor of teacher recruitment and retention (e.g., Renzulli et al., 2011; Wang et al., 2015). Moreover, job satisfaction is also linked to teachers’ occupational well-being and motivation, and some subscales are related to the disciplinary climate (Dicke et al., 2020). Teachers’ perceptions of the school climate have been found to be a key predictor of teacher self-efficacy, job satisfaction, and sense of stress (e.g., Borg, 1990; Hoy and Woolfolk, 1993; Kim and Loadman, 1994; Collie et al., 2012; Wilson et al., 2020). Warr (1999) and Harter et al. (2002) found that a high sense of workplace well-being can contribute to job retention.

Björnsson (2020) assessed the variations in Nordic teachers’ self-efficacy in multicultural classrooms using the Teaching and Learning International Survey (TALIS) 2018 data, in which the
influences of workload stress, teacher-student relations, job satisfaction, and disciplinary climate were also considered. The TALIS, which were organized by the Organization for Economic Co-operation and Development (OECD) from 2008 to 2018, assessed measures such as working conditions, beliefs, and attitudes from principals and teachers, with the goal to help countries review and develop policies that promote conditions for effective teaching and learning (Ainley and Carstens, 2018). Hereafter, response data from the third round of TALIS (i.e., 2018) are used.

Although research is emerging on teacher self-efficacy (TSE), job satisfaction (JS), school climate (SC), or workplace well-being and stress (WWS), it is critical to understand the internal relationship amongst them. However, the causality has not been taken into account. The primary questions of this article are: (1) to what extent do the extreme response styles vary across different countries/economies and (2) what variation is evident in causalities amongst TSE, JS, SC, and WWS across countries/economies? We propose the following hypotheses: H1. There exist extreme response styles for these four dimensions in TALIS 2018 data; H2. TSE, JS, and SC are positively correlated with each other, and negatively correlated with WWS; and H3. Distinct kernel causal paths exist across different countries/economies.

BACKGROUND

Teacher Self-Efficacy (TSE)

Self-efficacy was first defined by Bandura (1977) as belief in one's capability to accomplish desired outcomes, which can be grounded in mastery experience, vicarious experiences, social persuasion, and physical and emotional states. Similarly, TSE is defined as teachers' beliefs in their ability to solve challenges and difficulties accumulated within a teacher's professional career (Armor et al., 1976; Bandura, 1997; Tschanen-Moran and Hoy, 2001; Schwarzer and Hallum, 2008). Three dimensions of TSE are operationalized by the TALIS 2018 team (OECD, 2019, p. 285): self-efficacy in classroom management (SECLS), self-efficacy in instruction (SEINS), and self-efficacy in student engagement (SEENG).

Job Satisfaction (JS)

Job satisfaction can be defined as “a pleasurable or positive emotional state resulting from the appraisal of one’s job or job experiences” by Locke (1976, p. 1300), “the extent to which people like (satisfaction) or dislike (dissatisfaction) their jobs” by Spector (1997, p. 2), or “the state of mind determined by the extent to which the individual perceives her/his job-related needs to be being met” by Evans (1997, p. 833). Wang et al. (2018) conducted collaborative practices on teacher JS to examine how teachers perceive the comparison of actual job outcomes with desired ones.

In the TALIS 2018, three dimensions of JS are conceptualized and measured (OECD, 2019, p. 302): teacher JS with work environment (JSENV), job satisfaction with profession (JSPro), and satisfaction with target class autonomy (SATAT). The JSENV scale assesses the satisfaction of working at a specific school, the JSPro scale focuses on a global evaluation of the decision to become a teacher, and the SATAT scale measures the self-report of teaching at a specific class.

School Climate (SC)

The definition of SC is still an open question. SC can be referred to as the quality and character of school life depending on patterns of one’s personal experience (Cohen et al., 2009, p. 10). Zullig et al. (2010) found that social relationships in the school climate scale could be subdivided into three distinct areas: overall social environment, positive student–teacher relationships, and perceived exclusion/privelege. Among these areas, positive student-teacher relationships correlated positively with other school climate domains, and perceived exclusion/privelege correlated negatively with school connectedness (Zullig et al., 2010). Interested readers can refer to Zullig et al. (2010) and Thapa et al. (2013) for more school climate domains that are historically common.

The SC consisted of three subscales in the TALIS 2018 (OECD, 2019, p. 334): teachers' perceived disciplinary climate (DISC), teacher-student relations (STUD), and participation among stakeholders and teachers (STAKE). Among them, the DISC scale evaluates the class discipline, the STUD scale examines the self-report of the relationship between teachers and students, and the STAKE scale measures the distributed leadership.

Workplace Well-Being and Stress (WWS)

As a comprehensive social, physical, and emotional sense (Warr, 1990), employees' well-being not only matters to health and duty of care, but also links tangibly with effectiveness in the workplace (Lévi, 2000). In addition, workplace well-being can be treated as a fundamental element of successful organizations. Lazarus (1966) found that “stress arises when individuals perceive that they cannot adequately cope with the demands being made on them or with threats to their well-being.” Further, workplace stress has been defined by Colligan and Higgins (2006) as the variation of physical or/and mental response to an appraised challenge or threat posed by the workplace.

The WWS scale involved three subscales in the TALIS 2018 (OECD, 2019, p. 319): workplace well-being and stress (WELS), workload stress (WLOAD), and student behavior stress (STBEH). The WELS scale measures workplace well-being, stress, and its effect on other things; the WLOAD scale evaluate the stress connected to workload, and the STBEH scale evaluate the stress connected to classroom and student management.

Response Styles

Under the survey component of large-scale assessment, psychological constructs (e.g., beliefs and attitudes) are measured by rating or Likert-type scale self-reports. Baumgartner and Steenkamp (2001) found that items in such assessments are vulnerable to response styles (RS), or differences in how respondents tend to use the response options. Frequently used RS consist of extreme response style (ERS), midpoint response style (MRS), and acquiescent response style (ARS), which mean a tendency to choose extreme response options, to excessively
use the midpoint, and to agree with the item, respectively. In international studies, different response styles may be because of cultural variabilities (e.g., Hui and Triandis, 1989). For instance, Buckley (2009) considered heterogeneous response scales across countries for PISA 2006 data. Ju and Falk (2019) used a multilevel multidimensional nominal response model to accommodate extreme response styles. Let \( y_{ijg} \) denote the polytomous scored response of an examinee \( j \) \((j = 1, \ldots, f)\) from group \( g \) \((g = 1, \ldots, G)\) on item \( i \) \((i = 1, \ldots, I)\). The probability of endorsing category \( c \) \((c = 1, \ldots, C)\) is given by

\[
P(y_{ijg} = c | \theta_{ig}, a_{icg}, b_{icg}) = \frac{\exp(a_{icg} \theta_{ig} + b_{icg})}{\sum_{k=1}^{C} \exp(a_{ikg} \theta_{ig} + b_{ikg})} \tag{1}
\]

where \( \theta_{ig} \) denotes the latent trait (e.g., self-efficacy) of examinee \( j \) from group \( g \), \( a_{icg} \) is the discrimination parameter (i.e., slope) of item \( i \) for group \( g \), and \( b_{icg} \) is the intercept parameter of item \( i \) on category \( c \) for group \( g \). Moreover, many latent traits, such as self-efficacy (Woods, 2007b) and anxiety (Woods, 2006; Woods and Thissen, 2006), are not typically normally distributed in a population (Woods, 2015). Unsurprisingly, ERS may lead to non-normality of latent traits. Biased estimates of IRT model parameters will be obtained when the normal distribution assumption is violated (Woods, 2015). To handle the non-normal latent trait (LT) distribution, an empirical histogram (EH) method was proposed by Woods (2007a). The EH method is embedded in EM algorithm as a non-parametric approximation of the LT distribution that is simple to implement.

**Kernel Causality**

As pointed out by Zheng et al. (2012), Pearson’s correlation coefficient does not apply to asymmetric and/or nonlinear dependence, they proposed the generalized measures of correlation (GMC) to quantify the level of asymmetry in explaining variances. A pair of GMC can be expressed as

\[
\{ \text{GMC}(X|Y), \text{GMC}(Y|X) \} = \left[ 1 - \frac{E((Y - E(Y|X))^2)}{Var(Y)}, 1 - \frac{E((X - E(X|Y))^2)}{Var(X)} \right], \tag{2}
\]

where \( X \) and \( Y \) denote two variables, respectively.

A more refined version of the concept of Granger causality is kernel causality (Vinod, 2017), which can be treated as preliminary determination of causal directions among a set of variables. Kernel causality can be measured by GMC (Zheng et al., 2012). Let \( \delta = \text{GMC}(X \mid Y) - \text{GMC}(Y \mid X) \), then kernel cause is defined as follows (Vinod, 2017):

- if \( \delta < 0 \), i.e., \( \text{GMC}(Y \mid X) > \text{GMC}(X \mid Y) \), kernel cause = \( X \);
- if \( \delta = 0 \), i.e., \( \text{GMC}(Y \mid X) = \text{GMC}(X \mid Y) \), indeterminate cause;
- if \( \delta > 0 \), i.e., \( \text{GMC}(Y \mid X) < \text{GMC}(X \mid Y) \), kernel cause = \( Y \).

The hypothesis is \( H_0: \delta = 0 \) against \( H_1: \delta \neq 0 \), rejecting the null hypothesis (i.e., \( H_0 \)) suggests the existence of a statistically significant kernel causality. Vinod (2017) defined \( P(\text{cause}) \), which was calculated using the maximum entropy bootstrap algorithm, as the larger of the two rejection probabilities (i.e., reject \( \delta > 0 \) and \( \delta < 0 \)) in bootstrap resamples. A larger \( P(\text{cause}) \) means a larger rejection probability of \( H_0 \). When the relation between variables is linear and/or their joint distribution is close to normal, \( \delta \) is close to 0. In addition, 0.7 is recommended as the cut-off point to indicate a plausible kernel causality (Vinod, 2017).

**MATERIALS AND METHODS**

**Data Source and Sample**

In this article, we analyzed the response data of teachers in lower secondary schools (ISCED 2) from countries/economies which adopted the TALIS-PISA link option (i.e., Australia, Ciudad Autónoma de Buenos Aires (CABA) – Argentina, Colombia, Czechia, Denmark, Georgia, Malta, Turkey and Vietnam). We used listwise deletion for missing data. In total, the current study used data collected from 18,571 teachers at 1,512 schools (after initial data cleaning). The average length of teaching experience is 15.99 years with a 10.612 standard deviation, with 66.4% female teachers. **Table 1** summarizes sample configuration by country/economy, along with the country code used throughout the article.

**Measures**

Teacher self-efficacy and WWS are measured by questionnaires tailored to a 4-point Likert-type scale from “not at all” (1) to “a lot” (4), and the questionnaires of both JS and SC are tailored from “strongly disagree” (1) to “strongly agree” (4). The subtest lengths are 12, 13, 13, and 12 for TSE, JS, SC, and WWS, respectively. As a result, the corresponding total scores are 48, 52, 52, and 48 for each scale, respectively. Detailed information of items used in this study is presented in **Appendix Table 1**.

Responses on each scale are summarized in **Figure 1**; the darker the color is, the smaller value of the response option is. For TSE, JS, and WWS scales, more than 33% teachers chose extreme options. There were 2,072, 484, and 3 teachers answering the last

| Country/Economy Code | #Schools | #Teachers |
|----------------------|----------|-----------|
| Australia AUS        | 233      | 2,614     |
| CABA – Argentina ABA | 130      | 1,663     |
| Colombia COL         | 154      | 2,049     |
| Czechia CZE          | 219      | 3,028     |
| Denmark DNK          | 142      | 1,612     |
| Georgia GEO          | 191      | 1,704     |
| Malta MLT            | 55       | 1,048     |
| Turkey TUR           | 196      | 3,030     |
| Vietnam VNM          | 192      | 1,832     |
| **Total**            | 1,512    | 18,570    |

Source: OECD, TALIS 2018 Database.
options of all items for TSE, JS, and SC, respectively. For WWS, which can be treated as a negative sense, 133 teachers chose the first options of all items.

Analytic Procedure
The analytic procedure consists of two main stages. The first stage is to evaluate multi-group analysis to fit extreme response styles using item scores, and the second stage is to assess kernel causal inference using total scores.

Phase 1: Multi-Group Analysis
A multi-group analysis using the unidimensional nominal response model is conducted to fit response data with ERS. R package “mirt” (Chalmers, 2012) is used to fit the TALIS 2018 data. The EM algorithm with empirical histogram (Woods, 2007a) is used to estimate item parameters when latent trait distribution is non-normal or unknown, and the expected a posterior (EAP) method is used to estimate latent trait.

According to the TALIS 2018 technical report (OECD, 2019), each scale of the focused samples reached the metric invariance level, except STBEH subscale, which reached the configural invariance level. Therefore, the invariance = c('slopes') option is applied to support metric-level invariance and keep mean and variance of the population distribution consistent (i.e., mean = 0, variance = 1) for each group. Meanwhile, reliability testing is checked by Cronbach’s alpha reliability coefficient.

Phase 2: Kernel Casual Inference
R package “generalCorr” (Vinod, 2020) is used to calculate GMC and evaluate the direction of the kernel causal paths among TSE, SC, JS, and WWS. Total scores of each scale are used in this phase.

RESULTS
Multiple-Group Analysis
Table 2 shows the multi-group analysis results for the nine countries/economies, including Cronbach’s alpha reliability coefficient (α), the mean, standard deviation, skewness and excess kurtosis of total scores (i.e., μ, σ, βs and βk), the average and standard deviation of estimated latent traits (i.e., Û and σθ), and the average and confidence interval of estimated standard errors of latent traits (i.e., SE and CI).

All values of Cronbach’s alpha are larger than 0.6, which means an acceptable internal consistency. The Cronbach’s alpha of SC for each country/economy is the smallest, and some are less than an alternative acceptable cut-off point (i.e., 0.7). It appears that results from these data are reliable. The means and standard deviations of each scale’s total scores for each country/economy have similar trends to the case of total samples (shown in Figure 1). TSE of Colombia, JS and SC of Vietnam, and WWS of Georgia are the most centralized around their means.

The values of skewness of total samples are −0.291, −0.328, −0.061, and 0.363 for TSE, JS, SC, and WWS, respectively. And the values of excess kurtosis of total samples are −0.467, 0.224, 0.904, and −0.265 for TSE, SC, JS, and WWS, respectively. In other words, regarding total samples, TSE, JS, and SC are negatively skewed, and WWS is positively skewed. The distributions of TSE and WWS are platykurtic, and the distributions of JS and SC are leptokurtic. The skewness of Australia, CABA - Argentina, Malta, and Turkey are consistent with total samples, and the kurtosis of Australia and Vietnam are consistent with total samples.

In terms of LTs’ estimates (Table 2), Colombia’s teachers have the highest average level of TSE, JS, and SC, Georgia’s teachers have the highest average level of WWS, teachers with the lowest average level of TSE and JS are from the Czechia, teachers with the lowest average level of SC are from Vietnam, and those with the lowest average of WWS are from Malta. TSE estimates of Vietnam, JS estimates of Colombia, SC estimates of CABA – Argentina, and WWS estimates of Malta are the most centralized around the averages of their estimates. On the other hand, the standard errors of TSE for Vietnam, JS for Georgia, and SC and WWS for Malta are smallest. Note that the information of LTs’ estimates of SC for Vietnam presented in Table 2 is consistent as that shown in Figure 2.

Furthermore, histograms of latent traits measured each scale, which can be considered as an additional information to LT

2The bootstrap sampling number is set to be 50.

3The mean, standard deviation, skewness, and excess kurtosis are calculated by IBM SPSS statistics 26.
estimates, are shown in Figures 2–5. Most histograms for TSE, JS, and SC are asymmetric. This observation indicates that non-normal assumption is reasonable and necessary.

For TSE, Australia, the Czechia, Denmark, Malta, and Turkey have a wide range of estimates of LTs. For JS and WWS, Australia has the widest range of LTs' estimates. For SC, Georgia and Vietnam have a wide range of estimates.

Kernel Causal Inference

Tables 3–11 present the kernel causality measured by GMC and the Pearson’s correlation coefficients among these four scales for different countries/economies, respectively. The absolute values of GMC were larger than the corresponding Pearson’s correlation coefficient, and the trends of these statistics are similar. WWS is negatively correlated with other factors (i.e., TSE, JS, and SC) in most cases, excluding CABA – Argentina, the Czechia and Malta. For Australia, the Czechia, and Denmark, TSE and SC are also negatively correlated with a small correlation coefficient. The largest two correlation coefficients are that between JS and WWS and that between TSE and JS, and the smallest two correlation coefficients are that between TSE and SC and that between SC and WWS in most cases.

Comparing the pair of GMC, we can obtain the kernel cause. The probability of kernel causality is presented as P(cause) in the 8th column of Tables 3–11. Figure 6 summarizes the kernel causal paths for different countries/economies and total samples, which presents the kernel causality graphically. The solid line means kernel causality with acceptable probability [i.e., P(cause) ≥ 0.7], and the dotted line means unconvinced kernel causality [i.e., P(cause) < 0.7].
No kernel causality is always acceptable for all countries/economies. Comparing these countries/economies, most unconvinced kernel causalties appear when considering the relationship between WWS and other factors. For Australia, 3 of 6 kernel causal directions are identified with probability 1; 1 of 6, 3 of 6, 1 of 6, 4 of 6, and 2 of 6 kernel causal directions are determined with probability 1 for Columbia, the Czechia, Georgia, Malta, Turkey, and Vietnam, respectively. Only for Australia, all kernel causalities are acceptable with P(cause) larger than 0.88. As a result, the kernel causality for Turkey is the most stable, even though the probability of kernel causal direction from WWS to SC is 0.6. And the kernel causality among these
factors for Georgia is the least stable with 3 unconvinced kernel causal directions.

As a reference, Table 12 presents the GMC and corresponding information for total samples. For the total samples, the kernel causal direction from SC to TSE is unconvinced with probability 0.56. The directions for Australia, Denmark, and Vietnam are the same as those for total samples, but with different probabilities. Among these nine countries/economics and total samples, an explicit kernel causal factor⁴ exists in five countries/economics:

⁴The explicit kernel causal factor is defined as the factor which is always a kernel cause of other factors.
WWS for CBAB – Argentina, TSE for Colombia, JS for the Czechia, TSE for Georgia, and SC for Malta. Different kernel cause may result from different histories, cultures, and educational ecologies. On the other hand, a preferable kernel causal factor is recommended in four other countries/economics and total samples.

The preferable kernel causal factor is defined as the factor which is a kernel cause of more than half the factors (i.e., 2 of 3 factors in this research).

**DISCUSSION**

In the present study, teacher self-efficacy, job satisfaction, school climate, and workplace well-being and stress are shown to be related to teacher turnover (Borg, 1990; Warr, 1999; Tschannen-Moran and Hoy, 2001; Wang et al., 2015). Clarifying the relation among such factors can help researchers deeply explore the development path of high teaching quality and help policy makers formulate...
**FIGURE 5** | Histogram plots of latent traits (LTs) measured workplace well-being and stress (WWS).

**TABLE 3** | Summary of GMC for Australia.

| Code | Scenario | \( \rho \) | GMC(\( X \mid Y \)) | GMC(\( Y \mid X \)) | \( \delta \) | P(cause) | Kernel cause |
|------|----------|----------|-----------------|-----------------|--------|-----------|-------------|
| AUS  | TSE      | 0.2760   | 0.2816          | 0.2950          | -0.0134| 1         | TSE         |
|      | SC       | -.0512   | -.0677          | -.0834          | 0.0157 | 0.88      | SC          |
|      | WWS      | -0.0963  | -0.1080         | -0.1477         | 0.0397 | 1         | WWS         |
| JS   | SC       | 0.1816   | 0.1922          | 0.2008          | -0.0086| 0.96      | JS          |
|      | WWS      | -0.4871  | -0.4930         | -0.4922         | -0.0008| 0.88      | JS          |
| SC   | WWS      | -0.0664  | -0.1275         | -0.0810         | -0.0465| 1         | SC          |

P(cause), P(cause).
TABLE 4 | Summary of GMC for CABA – Argentina.

| Code | Scenario | ρ  | GMC(X| Y) | GMC(Y| X) | δ  | P(cause) | Kernel cause |
|------|----------|----|----------|----------|----|----------|-------------|
|      |          |    | X        | Y        |    |          |             |
| ABA  | TSE      | 0.3255 | 0.3358   | 0.3457   | −0.0099 | 0.88      | TSE         |
|      | SC       | 0.0546 | 0.0669   | 0.0895   | −0.0226 | 0.96      | TSE         |
|      | WWS      | −0.1143| −0.1366  | −0.1470  | 0.0104  | 0.92      | WWS         |
| JS   | SC       | 0.1811 | 0.1925   | 0.1887   | 0.0038  | 0.78      | SC          |
|      | WWS      | −0.3273| −0.3415  | −0.3427  | 0.0012  | 0.64      | WWS         |
| SC   | WWS      | 0.0836 | 0.1388   | 0.1011   | 0.0377  | 0.98      | WWS         |

P(cause), P(cause). Boldface entries are the values of P(cause) smaller than 0.7.

TABLE 5 | Summary of GMC for Colombia.

| Code | Scenario | ρ  | GMC(X| Y) | GMC(Y| X) | δ  | P(cause) | Kernel cause |
|------|----------|----|----------|----------|----|----------|-------------|
|      |          |    | X        | Y        |    |          |             |
| COL  | TSE      | 0.3301 | 0.3279   | 0.3487   | −0.0208 | 0.98      | TSE         |
|      | SC       | 0.0974 | 0.1600   | 0.1657   | −0.0057 | 0.82      | TSE         |
|      | WWS      | −0.2303| −0.2509  | −0.2343  | −0.0166 | 1         | TSE         |
| JS   | SC       | 0.2013 | 0.2606   | 0.2485   | 0.0121  | 0.88      | SC          |
|      | WWS      | −0.3778| −0.3972  | −0.3936  | −0.0036 | 0.96      | JS          |
| SC   | WWS      | −0.0792| −0.0989  | −0.1192  | 0.0203  | 0.64      | WWS         |

P(cause), P(cause). Boldface entries are the values of P(cause) smaller than 0.7.

TABLE 6 | Summary of GMC for the Czechia.

| Code | Scenario | ρ  | GMC(X| Y) | GMC(Y| X) | δ  | P(cause) | Kernel cause |
|------|----------|----|----------|----------|----|----------|-------------|
|      |          |    | X        | Y        |    |          |             |
| CZE  | TSE      | 0.2829 | 0.2989   | 0.2847   | 0.0142  | 1         | JS          |
|      | SC       | −0.0690| −0.1127  | −0.1321  | 0.0194  | 0.66      | SC          |
|      | WWS      | −0.1775| −0.1887  | −0.2067  | 0.018   | 0.82      | WWS         |
| JS   | SC       | 0.1560 | 0.1764   | 0.1961   | −0.0197 | 0.82      | JS          |
|      | WWS      | −0.4023| −0.4174  | −0.4064  | −0.0110 | 1         | JS          |
| SC   | WWS      | 0.0253 | 0.1295   | 0.0268   | 0.1027  | 1         | WWS         |

P(cause), P(cause). Boldface entries are the values of P(cause) smaller than 0.7.

TABLE 7 | Summary of GMC for Denmark.

| Code | Scenario | ρ  | GMC(X| Y) | GMC(Y| X) | δ  | P(cause) | Kernel cause |
|------|----------|----|----------|----------|----|----------|-------------|
|      |          |    | X        | Y        |    |          |             |
| DNK  | TSE      | 0.2293 | 0.2535   | 0.2894   | −0.0359 | 0.6       | TSE         |
|      | SC       | −0.0314| −0.0385  | −0.0933  | 0.0608  | 0.94      | SC          |
|      | WWS      | −0.1437| −0.1650  | −0.2040  | 0.0290  | 0.98      | WWS         |
| JS   | SC       | 0.2021 | 0.2296   | 0.2327   | −0.0031 | 0.82      | JS          |
|      | WWS      | −0.5192| −0.5293  | −0.5252  | −0.0041 | 0.76      | JS          |
| SC   | WWS      | −0.0545| −0.0895  | −0.0613  | −0.0282 | 0.66      | SC          |

P(cause), P(cause). Boldface entries are the values of P(cause) smaller than 0.7.

more efficient educational policy systems. In this article, we analyzed the TALIS-PISA linked data with extreme response style using the nominal response model with non-normal latent trait assumption. We compared them with multi-group analysis and explored the kernel causalities of each country/economy.

Comparison results of LTs are slightly different from those of total scores, and the non-normal LT assumption
TABLE 8 | Summary of GMC for Georgia.

| Code | Scenario | ρ  | GMC(X| Y) | GMC(Y| X) | δ  | P(cause) | Kernel cause |
|------|----------|----|---------|---------|----|-----------|--------------|
| X    | Y        |    |         |         |    |           |              |
| GEO  | TSE      | 0.4267 | 0.4299  | 0.4333  | −0.0034 | 0.76      | TSE          |
| SC   | JS       | 0.1965 | 0.2147  | 0.2568  | −0.0455 | 0.96      | TSE          |
| WWS  | −0.1483  | 0.3371 | 0.3468  | 0.4105  | −0.0637 | 1         | JS           |
| JS   | SC       | −0.2893 | 0.1101  | −0.1675 | 0.0209  | 0.66      | WWS          |

P(cause), P(cause). Boldface entries are the values of P(cause) smaller than 0.7.

TABLE 9 | Summary of GMC for Malta.

| Code | Scenario | ρ  | GMC(X| Y) | GMC(Y| X) | δ  | P(cause) | Kernel cause |
|------|----------|----|---------|---------|----|-----------|--------------|
| X    | Y        |    |         |         |    |           |              |
| MLT  | TSE      | 0.2396 | 0.2897  | 0.2524  | 0.0373  | 0.76      | JS           |
| SC   | JS       | 0.0686 | 0.0987  | 0.0848  | 0.0139  | 0.82      | SC           |
| WWS  | 0.0083   | 0.1857 | 0.1883  | 0.1882  | 0.0001  | 0.62      | TSE          |
| JS   | SC       | −0.4141 | 0.0004  | −0.0687 | −0.0807 | 0.72      | SC           |

P(cause), P(cause). Boldface entries are the values of P(cause) smaller than 0.7.

TABLE 10 | Summary of GMC for Turkey.

| Code | Scenario | ρ  | GMC(X| Y) | GMC(Y| X) | δ  | P(cause) | Kernel cause |
|------|----------|----|---------|---------|----|-----------|--------------|
| X    | Y        |    |         |         |    |           |              |
| TUR  | TSE      | 0.3609 | 0.3636  | 0.3802  | −0.0166 | 1         | TSE          |
| SC   | JS       | 0.1173 | 0.1329  | 0.1951  | −0.0622 | 1         | TSE          |
| WWS  | −0.1466  | 0.2481 | 0.2610  | 0.2890  | −0.0280 | 0.94      | JS           |
| JS   | SC       | −0.4285 | 0.0292  | −0.1004 | −0.1216 | 0.6       | WWS          |

P(cause), P(cause). Boldface entries are the values of P(cause) smaller than 0.7.

TABLE 11 | Summary of GMC for Vietnam.

| Code | Scenario | ρ  | GMC(X| Y) | GMC(Y| X) | δ  | P(cause) | Kernel cause |
|------|----------|----|---------|---------|----|-----------|--------------|
| X    | Y        |    |         |         |    |           |              |
| VNM  | TSE      | 0.2535 | 0.2636  | 0.2785  | −0.0149 | 0.9       | TSE          |
| SC   | JS       | 0.1469 | 0.1997  | 0.1911  | 0.0086  | 1         | SC           |
| WWS  | −0.0205  | 0.2110 | 0.2743  | 0.2860  | −0.0117 | 0.94      | JS           |
| JS   | SC       | −0.1564 | 0.2212  | −0.1721 | −0.0516 | 0.5      | JS           |

P(cause), P(cause). Boldface entries are the values of P(cause) smaller than 0.7.

is found to be reasonable. In terms of Pearson correlation coefficients and GMC, WWS is almost negatively correlated with other factors, and most other correlated relations are positive. Kernel cause is obtained by comparing the pair of GMC. Regarding the kernel causality, all causal directions for Australia are acceptable, of which the rejection probabilities
of $H_0$ are all larger than the cut-off point, but for other countries/economies, at least one causal direction is unconvinced (i.e., at least one rejection probability is smaller than the cut-off point). On the other hand, different explicit or preferable kernel causal factors are identified for different countries/economies, that may result from different histories, cultures, and educational ecologies.
These four factors analyzed in this article have influences on teacher's willingness of turnover. In the field of scientific research, this study adds to a growing body of research on relation among teacher's traits and educational large-scale survey (e.g., TALIS) analysis. The findings can inform teacher development and educational decision making.

To promote the professional and psychological development of teachers, some potentially effective interventions directed in an explicit or preferable kernel causal factor should be assessed, such as comprehensive training to help teachers develop a sense of self-efficacy (Burić and Moe, 2020; Fackler et al., 2021), a positive workplace-based program to promote teachers' job satisfaction (Ansley et al., 2019), or a customized mindfulness-based program to reduce teachers' stress and increase their wellbeing (Beshai et al., 2016).

For education policy makers, this study provides a novel direction to make policies for teachers with different purposes, as the trait may be changed by moderating its cause. It is worth noting that our results suggest that heterogeneous cultures and local characteristics lead to different kernel causal paths. Therefore, when making educational decision, strategies should be adjusted across different countries/economics.

**Limitations**

To the best of our knowledge, this is the first study to assess kernel causality of teachers' traits (i.e., self-efficacy, job satisfaction, school climate, and workplace well-being and stress in this article) using GMC. While the results are informative, this study can be extended in a number of directions. First, the mediation effects of these factors can be evaluated and compared. Second, as the TALIS-PISA linked data are available, the relation among such factors can also be investigated from the perspective of principals and/or students. The conclusion obtained through different angles of view will be more objective with more credibility. The relation between teaching quality and students' achievement or students' well-being (Beshai et al., 2016).

Furthermore, we can use the multilevel or and multidimensional nominal response model to analyze ERS (Ju and Falk, 2019). Finally, international comparative study of educational policies and education ecological environment should be conducted to explain the different causal paths among such factors.

**CONCLUSION**

The assessment of kernel causality indicates the explicitly preferable kernel causal factor among these four factors. Further, the results of the multi-group analysis discussed in this article support the hypothesis that there exist extreme response styles and the rationale to adopt non-normal assumption of latent traits' distribution. These findings contribute to the literature on quantitative research of causality, beyond the existing knowledge based on correlation or association. In addition, our study has identified important new areas to be considered when exploring the relationship among teachers' and students' traits under educational settings.

**DATA AVAILABILITY STATEMENT**

Publicly available datasets were analyzed in this study. This data can be found here: https://www.oecd.org/education/talis/talis-2018-data.htm OECD, TALIS.

**AUTHOR CONTRIBUTIONS**

XZ provided original thoughts and key technical support and completed the writing of the article. CZ did the data analysis. YX reviewed the literature. SL and ZW provided key theoretical support. All authors contributed to the article and approved the submitted version.

**FUNDING**

This study was funded by Grant 20YJC880124 from the Humanities and Social Sciences Youth Foundation, Ministry of Education of the People's Republic of China.
REFERENCES

Ansley, B. M., Houchins, D., and Varjas, K. (2019). Cultivating positive work environments. Teach. Teach. Educ. 99:103255. doi: 10.1016/j.tate.2020.10.3255

Granizer, H., and Perera, H. N. (2019). Relations among teachers’ self-efficacy beliefs, engagement, and work satisfaction: a social cognitive view. Contemp. Educ. Psychol. 58, 75–84. doi: 10.1016/j.cedpsych.2019.02.003

Harter, J. K., Schmidt, F. L., and Keys, C. L. (2002). “Well-being in the workplace and its relationship to business outcomes: a review of the Gallup studies,” in Flourishing: The Positive Person and the Good Life, eds C. L. Keys and J. Haidt (Washington, DC: American Psychological Association), 205–224. doi: 10.1037/10594-009

Hoy, W. K., and Woolfolk, A. E. (1993). Teachers’ sense of efficacy and the organizational health of schools. Elem. Sch. J. 93, 355–372. doi: 10.1086/461729

Hui, C. H., and Triandis, H. C. (1989). Effects of culture and response format on extreme response style. J. Cross Cult. Psychol. 20, 296–309. doi: 10.1177/0022441892030004

Ju, U., and Falk, C. F. (2019). Modeling response styles in cross-country self-reports: an application of a multilevel multidimensional Nominal Response Model. J. Educ. Meas. 56, 169–191. doi: 10.1111/jedm.12205

Kim, I., and Loadman, W. E. (1994). Predicting Teacher Job Satisfaction. Columbus, OH: Ohio State University.

Lazarus, R. S. (1966). Psychological Stress and the Coping Process. New York, NY: McGraw-Hill.

Lévi, L. (2000). Stressors at the workplace: theoretical models. Occup. Med. 15, 69–106

Liu, S., and Hallinger, P. (2018). Principal instructional leadership, teacher self-efficacy, and teacher professional learning in China: testing a mediated-effects model. Educ. Adm. Q. 54, 501–528. doi: 10.1177/0013838X17869048

Locke, E. A. (1976). “The nature and causes of job satisfaction,” in Handbook of Industrial and Organizational Psychology, ed. M. D. Dunnette (Chicago, IL: Rand McNally), 1297–1343.

Møk, A., and Katz, I. (2020a). Emotion regulation and need satisfaction shape a motivating teaching style. Teach. Teach. 1–18. doi: 10.1080/15340602.2020.1777960

Møk, A., and Katz, I. (2020b). Self-compassionate teachers are more autonomy supportive and structuring whereas self-derogating teachers are more controlling and chaotic: the mediating role of need satisfaction and burnout. Teach. Teach. Educ. 96:103173. doi: 10.1016/j.tate.2020.10.3173

OECD (2019). TALIS 2018 Technical Report. Paris: OECD Publishing.

Renzulli, L. A., Parrott, H. M., and Beattie, I. R. (2011). Racial mismatch and school type teacher satisfaction and retention in charter and traditional public schools. Sociol. Educ. 84, 23–48. doi: 10.1177/0038040710315072

Rev., W. J. (2013). Instructional Leadership Practices and Teacher Efficacy Beliefs: Cross-National Evidence from TALIS. Unpublished doctoral dissertation. Tallahassee, FL: Florida State University.

Schwarzer, R., and Hallum, S. (2008). Perceived teacher self-efficacy as a predictor of job stress and burnout: mediation analyses. Appl. Psychol. 57, 152–171. doi: 10.1111/j.1464-0597.2008.00359.x

Smylie, M. A. (1988). The enhancement function of staff development: organizational and psychological antecedents to individual teacher change. Am. Educ. Res. J. 25, 1–30. doi: 10.3102/002196241250010001

Spector, P. E. (1997). Job Satisfaction: Application, Assessment, Causes, and Consequences. London: Sage.

Thapa, A., Cohen, J., Guffey, S., and Higgins-D’Alessandro, A. (2013). A review of school climate research. Rev. Educ. Res. 83, 357–385. doi: 10.3102/0034654313483907

Tschanen-Moran, M., and Hoy, A. W. (2001). Teacher efficacy: capturing an elusive construct. Teach. Teach. Educ. 17, 783–805. doi: 10.1016/0742-051x(01)00036-1

Vinod, H. D. (2017). Generalized correlation and kernel causality with applications in development economics. Commun. Stat. Simul. Comput. 46, 4513–4534. doi: 10.1080/03610918.2015.1122048

Vinod, H. D. (2020). generalCorr: Generalized Correlations and Initial Causal Path. R package version 1.1.8. New York, NY: Fordham University.

Wang, H., Hall, N. C., and Rahimi, S. (2015). Self-efficacy and causal attributions in teachers: effects on burnout, job satisfaction, illness, and...
Zhang et al. Kernel Causality in TALIS quitting intentions. Teach. Teach. Educ. 47, 120–130. doi: 10.1016/j.tate.2014.12.005

Wang, K., Chen, Z., Luo, W., Li, Y., and Waxman, H. (2018). Examining the differences between the job satisfaction of STEM and non-STEM novice teachers with leaving intentions. EURASIA J. Math. Sci. Technol. Educ. 14, 2329–2341. doi: 10.29333/ejmst/89516

Warr, P. (1990). The measurement of well-being and other aspects of mental health. J. Occup. Psychol. 63, 193–210. doi: 10.1111/j.2044-8325.1990.tb00521.x

Warr, P. (1999). “Well-being and the workplace,” in Well-Being: The Foundations of Hedonic Psychology, eds D. Kahneman, E. Deiner, and N. Schwarz (New York, NY: Russell Sage), 392–412.

Wilson, C., Marks Woolfson, L., and Durkin, K. (2020). School environment and mastery experience as predictors of teachers’ self-efficacy beliefs towards inclusive teaching. Int. J. Incl. Educ. 24, 218–234. doi: 10.1080/13603116.2018.1455901

Woods, C. M. (2006). Ramsay-curve item response theory (RC-IRT) to detect and correct for nonnormal latent variables. Psychol. Methods 11, 253–270. doi: 10.1037/1082-989X.11.3.253

Woods, C. M. (2007a). Empirical histograms in item response theory with ordinal data. Educ. Psychol. Meas. 67, 73–87. doi: 10.1177/0013164406288163

Woods, C. M. (2007b). Ramsay curve IRT for Likert-type data. Appl. Psychol. Meas. 31, 195–212. doi: 10.1177/0146621606291567

Woods, C. M. (2015). “Estimating the latent density in unidimensional IRT to permit non-normality,” in Handbook of Item Response Theory Modeling: Applications to Typical Performance Assessment, eds S. P. Reise and D. A. Revicki (New York, NY: Routledge), 60–84.

Woods, C. M., and Thissen, D. (2006). Item response theory with estimation of the latent population distribution using spline-based densities. Psychometrika 71:281. doi: 10.1007/s11336-004-1175-8

Zheng, S., Shi, N. Z., and Zhang, Z. (2012). Generalized measures of correlation for asymmetry, nonlinearity, and beyond. J. Am. Stat. Assoc. 107, 1239–1252. doi: 10.1080/01621459.2012.710509

Zullig, K. J., Koopman, T. M., Patton, J. M., and Ubbes, V. A. (2010). School climate: historical review, instrument development, and school assessment. J. Psychoeduc. Assess. 28, 139–152. doi: 10.1177/0734282909344205

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Zhang, Zhao, Xu, Liu and Wu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.
### APPENDIX

#### APPENDIX TABLE 1 | Items used in this study from TALIS 2018.

| Scale | Subscale | Items | Original TALIS code |
|-------|----------|-------|---------------------|
| TSE   | SECLS    | TT3G34D, TT3G34F, TT3G34H, TT3G34I | T3SECLS |
|       | SEINS    | TT3G34C, TT3G34J, TT3G34K, TT3G34L | T3SEINS |
|       | SEENG    | TT3G34A, TT3G34B, TT3G34E, TT3G34G | T3SEENG |
| JS    | JSENV    | TT3G53C*, TT3G53E, TT3G53G, TT3G53J | T3JSENV |
|       | JS PRO   | TT3G53A, TT3G53B, TT3G53D*, TT3G53F* | T3JS PRO |
| SC    | DISC     | TT3G41A, TT3G41B*, TT3G41C, TT3G41D | T3DISC |
|       | STUD     | TT3G49A, TT3G49B, TT3G49C, TT3G49D | T3STUD |
|       | STAKE    | TT3G48A, TT3G48B, TT3G48C, TT3G48D, TT3G48E | T3STAKE |
| WWS   | WELS     | TT3G51A, TT3G51B*, TT3G51C, TT3G51D | T3WELS |
|       | WLOAD    | TT3G52A, TT3G52B, TT3G52E, TT3G52C, TT3G52D | T3WLOAD |
|       | STBEH    | TT3G52F, TT3G52G, TT3G52H | T3STBEH |

TSE, teacher self-efficacy; SECLS, self-efficacy in classroom management; SEINS, self-efficacy in instruction; SEENG, self-efficacy in student engagement; JS, job satisfaction; JSENV, job satisfaction with work environment; JS PRO, job satisfaction with profession; SATAT, satisfaction with target class autonomy; SC, school climate; DISC, teachers' perceived disciplinary climate; STUD, teacher-student relations; STAKE, participation among stakeholders, teachers; WWS, workplace well-being and stress; WELS, workplace well-being and stress; WLOAD, workload stress; STBEH, student behavior stress.

*Denotes the reverse coded item.

Source: OECD, TALIS 2018 Database.