1969

Initial development and validation of the multidimensional teacher victimization scale among Chinese teachers

Chunyan Yang, Stephanie Secord-Fredrick, Amanda Nickerson, Lyndsay Jenkins and Jia-Shu Xie

Initial Development and Validation of the Multidimensional Teacher Victimization Scale among Chinese Teachers

Chunyan Yang[1]

Stephanie Secord Fredrick [2]

Amanda Nickerson [3]

Lyndsay Jenkins [4]

Jia-Shu Xie[5]

[1] Department of Counseling, Clinical, and School Psychology, University of California - Santa Barbara, Santa Barbara, USA

[2] Department of Psychology, Central Michigan University, Mt Pleasant, MI, USA

[3] Alberti Center for Bullying Abuse Prevention, Department of Counseling, School, and Educational Psychology, University at Buffalo, The State University of New York, Buffalo, NY, USA

[4] College of Education, Florida State University, Tallahassee, Florida, USA

[5] School of Education, Hunan Normal University, Changsha, Hunan Province, China

Corresponding Author: Chunyan Yang

Email: cyang@education.ucsb.edu

Postal Address: Gevirtz School of Education, University of California - Santa Barbara, Santa Barbara, USA 93106-9490

Abstract

The Multidimensional Teacher Victimization Scale (MTVS) was developed to provide schools with a comprehensive and psychometrically sound self-report scale to assess teachers' perception of teacher-targeted aggressive and violent behaviors perpetrated by students. Confirmatory factor analyses, conducted on a sample of 1,711 teachers (7th to 12th grade) from 58 schools in China, showed that a second-order model consisting of one general Teacher Victimization (TV) factor and six lower-order factors (i.e., Physical TV, Social TV, Verbal TV, Cyber TV, Sexual Harassment, and Personal Property Offenses) best represented the data. Measurement invariance tests showed that the scale's factor structure was consistent across middle and high schools and across gender. Latent mean differences were found among gender, with male and female teachers reporting different levels in different types of victimization. Latent mean differences indicated that similar levels of teacher victimization were reported across middle school and high school teachers. As evidence of the scale's concurrent validity, the total teacher victimization score correlated significantly with students' self-reported total burnout score and three subscale scores, including emotional exhaustion, accomplishment, and depersonalization. Adequate internal consistency reliability and test-retest reliability were achieved in both the total scale and subscales. Implications for using MTVS as a school-wide violence assessment tool for use in Chinese schools and other cultural groups are discussed.

Keywords: Teacher victimization; measurement invariance; psychometric property

Impact and Implication Statement: The findings provide empirical supports for the validity and reliability of the Multidimensional Teacher Victimization Scale, that could be used in school-based programs targeting school violence and school safety.

Initial Development and Validation of the Multidimensional Teacher Victimization Scale among Chinese Teachers

Despite worldwide attention to school violence against students, research on teacher-directed violence perpetrated by students has emerged only recently as a critical issue. Teachers' victimization experiences are associated with a wide range of negative outcomes, such as increased physical and psychological symptoms (Bounds & Jenkins, 2016; Wilson, Douglas, & Lyon, 2011), higher rates of burnout, stress, and teacher turnover (Ingersoll, 2011; Ozkilic & Kartal, 2012), and decreased sense of school safety, morale, and job satisfaction (Wilson et al., 2011). Challenges remain for researchers, educators, and policy makers to accurately understand the prevalence and severity of teacher victimization due to the scarcity of psychometrically-sound multi-dimensional measures (e.g., Dzuka & Dalbert, 2007; Espelage et al., 2013). Moreover, the majority of studies on teacher victimization (TV) and teacher-directed violence have been conducted in the U.S. and Europe. There is a dearth of valid measures of TV used in non-European and developing countries, such as China. To address this, the current study aims to develop the Multidimensional Teacher Victimization Scale (MTVS) to assess Chinese teachers' perceptions of teacher-directed violent and aggressive behaviors from students in middle and high schools. Specifically, we used confirmatory factor analysis to examine its factor structure and measurement invariance across gender and grade levels; we also examined its concurrent validity, internal consistency, and test-retest reliability.

Limitations of Existing Teacher Victimization Measures

In the relatively small but growing body of literature, TV has been measured by instruments lacking sound psychometric support. In a recent systematic review of 37 published studies related to TV, Reddy and colleagues (2018) found that 24 studies used author-developed

questionnaires or surveys to assess TV experiences. Moreover, 52% of the studies reviewed did not provide evidence of reliability and validity for the TV measure, and internal consistency reliability was the only type of psychometric evidence for the remaining 48% of studies. The main limitations of the existing TV measures are discussed below.

Lack of multidimensional measures. Decades of research on peer victimization in youth has recognized that victimization is a multi-dimensional construct (Marsh et al., 2011; Mynard & Joseph, 2000). Although consensus has not emerged, verbal (i.e., use of aggressive words), physical (i.e., actual or threatened bodily contact or physical injury), and relational/social victimization (i.e., manipulating or damaging social relationships or standing) are the three most commonly recognized dimensions. With the increased use of technology and social media by youth over the past 15 years, cyber victimization has been recognized as a form of aggression which occurs via digital means (i.e., cell phones, computers, or tablets) on platforms such as texting or social media (Gardella, Fisher, & Teurbe-Tolon, 2017; Kowalski, Giumetti, Schroeder, & Lattanner, 2014). Moreover, personal property offenses, which include theft of personal property, damage to personal property, or attacks on property, is another commonly identified form of peer victimization (Mynard & Joseph, 2000). Sexual harassment is a widely studied form of violence experienced by not only students but also teachers (McMahon et al., 2014).

Consistent with the peer victimization literature, these six forms of victimization (i.e., physical, verbal, relational, and cyber victimization, sexual harassment, and personal property offenses) have been directly or indirectly measured in a few existing multi-dimensional TV scales (Kauppi & Pörhölä, 2012; Mooij, 2011). For example, Kauppi and Pörhölä (2012) developed a 22-item scale measuring Finnish teachers' experience of being subjected to direct verbal bullying, direct nonverbal bullying, physical bullying, indirect private bullying, and

indirect public bullying by students in primary and lower secondary schools. This measure has low internal consistency coefficients, with subscale alpha coefficients ranging from .38 to .74. An Internet-based questionnaire was used to measure teacher-directed violence across verbal, social, mild physical, severe physical, material, and sexual domains in Dutch secondary schools (Mooij, 2011). Although this scale is comprehensive, only descriptive statistics of the frequency of each form of violent behavior was examined, and no evidence about its reliability and validity was reported.

Lack of measurement invariance support. Existing studies have reported that teachers' victimization experiences vary across subgroups (i.e., gender, race/ethnicity, and grade levels). For example, male teachers are more likely than female teachers to experience threats, physical violence, and multiple forms of aggression from students (Berg & Cornell, 2016; Martinez et al., 2015; McMahon et al., 2014), while female teachers were more likely to experience verbal and nonphysical forms of violence, such as property damage (Berg & Cornell, 2016; Moon, Morash, Jang, & Jeong, 2015; Wei et al., 2013). Also, white teachers reported higher rates of victimization compared to African American and Latino teachers (Dworkin et al., 1988; Martinez et al., 2015). Moreover, secondary school teachers are more likely to experience violence than those who teach earlier grades (Kauppi & Pörhöla, 2012; Lyon & Douglas, 1999). However, none of the above studies have conducted rigorous measurement invariance tests to establish that group differences are real and not simply artifacts of measurement error.

Lack of psychometrically-sound teacher victimization measures in China. With the rapid social, cultural, and economic transformations in China in the past decades, violence against teachers has gained media attention (Dong, 2010; Guthrie, 2012; Wang, 2014). However, empirical research on TV is very limited, largely due to the lack of psychometrically-sound

Chinese TV measures. In the existing literature, we found only two studies using Likert-scales to measure teacher-targeted violent and aggressive behavior perpetrated by students (Chen & Astor, 2011; Dong, 2010). Dong (2010) used a four-dimension scale to measure students' and teachers' perceptions of verbal, behavioral, relational, physical, and cyber victimization experienced by teachers, but its factor structure was not empirically supported using factor analysis. Chen and Astor (2011) used a 7-item scale to measure physical violence (2 items), verbal violence/threat (3 items), and emotional violence/harassment (2 items) against teachers. Although the scale's factor structure was examined, many items had low factor loadings, the overall scale's reliability was low, and the reliability coefficients for subscales were not reported.

To address the above limitations, a Multidimensional Teacher Victimization Scale (MTVS) was developed among a large and geographically diverse group of teachers in Mainland China to access Chinese teachers' overall experiences with aggressive and violent behaviors. It was developed to measure overall teacher victimization and its six subtypes, including physical, verbal, social/relational, cyber victimization, sexual harassment, and personal property offenses. Confirmatory factor analysis was used to examine and compare the hypothesized and alternative factor structures of MTVS. Based on the final model, its measurement invariance and latent mean differences were examined across gender and grade levels. Concurrent validity was also tested by examining its associations with teacher burnout; the scale's internal consistency and test-retest reliability were also examined.

Method

Participants

The study consisted of 1,711 teachers (7^{th} to 12^{th} grade) from 58 schools across eight provinces in mainland China. The sample included 893 teachers from 33 middle schools, 505 teachers from 12 high schools, and 313 teachers from 12 schools with combined middle and high

school levels. Online Supplemental Table 1 includes demographic information about the teachers and schools.

Data Collection Procedure

Data collection was conducted by a research team from a university focusing on training K-12 teachers in China. During the summer of 2017, 58 schools were selected through the alumni network with the aim to form a representative sample of Chinese middle and high schools. Although the 58 schools comprised a sample of convenience and were not selected randomly, schools' location (urban and non-urban), type (private and public), and academic achievement level (i.e., lower-end, middle, and higher-end levels) were taken into consideration. At the beginning of the 2017-2018 school year, participating schools were provided with the option of the web-based version or the paper version of the survey. For schools choosing the web-based survey, a message with a link to the survey and detailed information about the purpose and instructions of the study was sent to all teachers. For schools choosing the paper version of the survey, the research team provided a designated person in each school (e.g., school administrator, lead teacher, school counselor, or graduate student) with a 45-minute online training about the purpose of the study, random sampling requirements, and data collection procedures, including protections for teachers' confidentiality. This person administered the surveys and mailed them to the local Chinese university. No incentives were provided for schools' participation, other than learning of the prevalence of TV in their schools. Schools' survey response rates ranged from 4% to 71% (M = 24.31%, median = 73.03%), number of respondents ranged from 9 to 101 (M = 29, median = 25). Missing responses to individual survey items ranged from 0.6% to 3.4%. Missing responses to composite scores ranged from 0.9%

(Verbal TV subscale) to 3.7% (Sexual Harassment subscale). All measures and administration procedures were approved by the researchers' university Institutional Review Board.

Scale Development and Preliminary Analyses

To develop the MTVS, a team of three bilingual school-based researchers first drafted a list of items based on a thorough review of the existing measures of TV and peer victimization in both the Chinese and English literature. For each of the six main forms of TV, 5 to 8 items were developed. The research team then discussed the initial set of items with a group of teachers and administrators who worked in the Chinese schools and asked them to evaluate the face validity and readability of the measure. Based on feedback from the teacher groups, items were modified and refined, and 32 items were chosen to create the initial version of MTVS, which included six subscales: physical TV (6 items), verbal TV (5 items), relational TV (5 items), cyber TV (5 items), sexual harassment (6 items), and personal property offenses (5 items). Based on data collected using the 32-item MTVS, confirmatory factor analysis (CFA) was first conducted with the full sample to examine if the scale was supported by the hypothesized 6-factor second-order model with one general teacher-victimization factor and six lower-order factors. Results showed that two items had very low loadings (below .40) on the hypothesized factor or dual loadings on other factors. Thus, these two items were eliminated to improve the model fit. Based on the 30-item MTVS, confirmatory factor analysis was conducted in different gender and grade-level subgroups. Six items with very lower factor loading or high dual-loading on other subscale were eliminated to achieve adequate model fit in each subgroup. After the preliminary CFA analyses, 24 items were left to form the refined version of the MTVS for comprehensive validity and reliability analyses.

Measures

Multidimensional Teacher Victimization Scale (MTVS). The refined version of Multidimensional Teacher Victimization Scale consists of 24 items that assess individual teachers' perceptions of how often they have been victims of six forms of violent behavior perpetrated by students "during the last school year" on a 5-point Likert scale (1 = *Never*, 2 = *Once in several months*, 3 = *Several times in one month*, 4 = *Once a week, 5= More than once a week*). Cronbach's alpha reliability coefficients were .99 for Physical TV, .92 for Social TV, .91 for Verbal TV, .97 for Cyber TV, .87 for Sexual Harassment, and .96 for Personal Property Offenses.

Chinese Teachers' Job Burnout Questionnaire-Revised (CTJBQ-R; Wu, Qi, Yu, & Zhang, 2016**).** The CTJBQ-R measures Chinese primary and secondary school teachers' job burnout perceptions with 23 items on a 7-point Likert scale (0 = *Never,* 1 = *Rarely*, 2 = *Occasionally*, 3 = *Sometimes*, 4 = *Often*, 5 = *Quite Often*, and 6 = *Always*). The CTJBQ-R was adapted from the Burnout Inventory-Educators Survey (MBI-ES) and its validity and reliability have been examined among Chinese teachers (Wang, Liu, & Wu, 2003; Wu, Qi, Yu, & Zhang, 2016). In the present study, CFA results demonstrated that the CTJBQ-R was best represented by a three-factor correlation model consisting of Emotional Exhaustion (8 items; $\alpha = .92$), Accomplishment (9 items; $\alpha = .86$), and Depersonalization (6 items; $\alpha = 75$), $\chi^2 = 834.25$ (206, N =806), $p < .001$; CFI = .916, RMSEA = .061, SRMR = .062. The mean composite scores of each of the three subscales were used to examine the concurrent validity of MTVS.

Data Analyses of the 24-item Version of MTVS

The analysis on the 24-item version of MTVS was conducted in four stages. The first two stages of analysis were conducted in Mplus 7.31 (Muthén & Muthén 1998-2015) utilizing the full information maximum likelihood (FIML) estimator. Prior to conducting the CFA in stage 1,

Intraclass Correlation (ICC) and design effect (DEEF) were first calculated to examine if clustering in the data needed to be taken into account. Results showed that the DEEF was less than 2.0 (DEEF = 1.87); thus, data were modeled without the consideration of its nesting structure (Muthén & Satorra, 1995).

In the first stage, CFA was conducted to determine the final model for the MTVS. CFA is considered the method of choice when the factorial structure of a scale is hypothesized (Thompson, 2004). A hypothesized second-order factor model was first tested. Three alternative models (a one-factor model, a six-factor model, and a bifactor model) were then estimated and compared with the hypothesized second-order factor model. During analysis in the first stage, the sample was randomly divided into two subsamples for cross validation purposes. One half of the sample was used to analyze and compare model fit for the hypothesized model and the three alternative models. The final model was then replicated using the second half of the sample. When evaluating the hypothesized and alternative models, model fit was assessed using Satorra–Bentler scaled chi-square values ($\Delta S\text{-}B\chi^2$), Akaike Information Criterion (AIC) values, and three other fit indices: Comparative Fit Index (CFI), Root Mean-Square Error of Approximation (RMSEA), and Standardized Root Mean-Square Residual (SRMR). The $\Delta S\text{-}B\chi^2$ test was used to compare nested models (Asparouhov & Muthén, 2010) and AIC values were used to compare non-nested models. It is desirable to have CFI values > .90 and RMSEA and SRMR values < .08 (Hu & Bentler, 1998). The second-order model was selected as the final model from the first stage.

In the second stage, measurement invariance was tested in a hierarchical sequence with increasingly restrictive steps to investigate whether the factor structure of the final model was statistically equivalent across gender (i.e., male and female teachers) and grade level (i.e., middle

and high school level teachers). We followed five steps suggested by Chen, Sousa, and West (2005): (a) configural invariance; (b) first-order factor loading invariance; (c) first- and second-order factor loading invariance; (d) first- and second-order factor loading and intercepts of measured variables invariance; and (e) first- and second-order factor loadings, and intercepts of measured variables and first-order factors invariance. Each pair of models in the sequence is nested because a set of parameters are constrained to be equal across groups in the more restricted model. To compare the fit for two nested models, $\Delta S\text{-}B\chi^2$ (Asparouhov & Muthén, 2010) and the goodness-of-fit indexes (Cheung & Rensvold, 2002) were used. However, because the performance of the chi-square difference test is also affected by non-normality and large sample size, we followed the recommendation by Cheung and Rensvold (2002) and considered a difference of larger than .01 in the change of CFI as an indication of a meaningful change in model fit for testing measurement invariance.

The third stage examined latent mean differences of TVs across gender and grade levels. The male and middle school groups were chosen as the reference group(s), with the latent mean set to zero. The latent means of the female and high school group were freely estimated, with values reflecting latent mean differences across gender and grade levels. Statistical significance of latent mean differences was determined using the z-statistic (Aiken, Stein, & Bentler, 1994). Effect sizes and confidence intervals associated with the latent mean differences were estimated according to the guidelines of Hancock (2001).

In the fourth stage, we examined the internal consistency reliability of each subscale. Using the paired data collected from 61 teachers, 3 months following the initial data collection, we also examined the test-retest reliability of each subscale. Moreover, we examined the evidence of concurrent validity by examining the correlations between MTVS total scale and

subscale scores and teacher-reported burnout total scale and subscale scores. In addition, correlations among the subscales and the full scale of MTVS were computed to examine their relative independence and the degree to which they assessed the TV construct.

Results

Stage 1: Confirmatory Factor Analysis (CFA)

Based on the first randomly selected half of the sample, adequate model fits were achieved in both the hypothesized second-order model (χ^2 = 608.95 (*df* = 246, *N* = 846), $p <$ .001; CFI = .917, RMSEA = .042, 90% CI = [.038, .046], and SRMR = .035) and four-factor model (χ^2 = 574.73 (*df* = 246, *N* = 846), $p <$ .001; CFI = .922, RMSEA = .041, 90% CI = [.037, .045], and SRMR = .032), whereas the one-factor model , the most parsimonious of the three alternative models, yielded poor fit statistics (χ^2 = 1356.26 (*df* = 252, *N* = 846), $p <$ .001; CFI = .775, RMSEA = .072, 90% CI = [.068, .076], and SRMR = .038). The bifactor model failed to converge. When the four-factor model was compared to the proposed nested second-order model, the magnitude of the model fit differences was small ($\Delta S\text{-}B\chi^2$= 29.62 (Δ *df* = 9), p = 0.00) and the differences of the model fit indices (i.e., CFI, SRMR, and RMSEA) were negligible (ranging from 0.001 to 0.004). Considering that the second-order model is more parsimonious and it is also consistent with theoretical conceptualizations of TV, the second-order model with one higher-order factor and six lower-order factors (i.e., physical, verbal, social, cyber victimization, sexual harassment, and personal property offenses) was selected as the final model.

Robust support for the second-order model was found when the CFA was replicated with the second randomly selected half of the sample, χ^2 = 608.30 (*df* = 246, *N* = 842), $p <$ .001; CFI = .925, RMSEA = .057, 90% CI = [.037, .046], and SRMR = .038. Items generally had similar

standardized factor loadings in the two halves of the sample (see Supplemental Table 2). Because no appreciable differences in the fit indices or factor loadings were found, all subsequent analyses were run with the full sample. The standardized factor loadings for the second-order model with the full sample size are illustrated in Figure 1. Fit statistics for the second-order model in the full sample and subgroups by gender and grade levels indicated robust support for the second-order model across male and female teachers and across middle and high school teachers (see Online Supplemental Table 3).

Stage 2: Measurement Invariance Test

Following the procedure described previously, measurement invariance of the second-order model was tested in a hierarchical sequence with five increasingly restrictive steps to examine whether the factor structure of the second-order model was statistically equivalent across gender and grade level. As shown in Online Supplemental Table 4, the model testing configural invariance (Models 1) had adequate fit across both gender and grade levels; the comparison of model fit between Models 1 vs. 2, Models 2 vs. 3, Models 3 vs. 4, and Models 4 vs. 5 suggested that all five levels of measurement invariance were achieved across gender and grade levels.

Stage 3: Latent Mean Differences of TVs across Gender and Grade Levels

When gender differences in latent means were compared, male teachers reported significantly more frequent total TV, Social TV, Verbal TV, Cyber TV and Personal Property Offences, but not Physical TV and Sexual Harassment. There were no significant grade-level differences across total TV scale scores and subscale scores (see Table 1).

Stage 4: Reliability and Concurrent Validity of MTVS

As shown in Online Supplemental Table 5, Cronbach's alpha reliability coefficients of the

total scale and six subscales ranged from .96 to .98 for the full sample. Among the full sample, 61 teachers participated the retest 3 months later and the test-retest reliability was .72 for the TV total scale and ranged from .56 (Physical TV) to .77 (personal property offenses) across the six subscales.

The concurrent validity of the MTVS was examined using total TV scores' correlations with three subscale scores of CTJBQ-R in the full sample. Results in Table 2 indicated that total TV was significantly correlated with all three subtypes of burnout in the expected directions (all *p*s < .001). When the association between each type of TV and burnout were examined, all six types of TVs had stronger association with depersonalization than emotional exhaustion and accomplishment. In particular, Verbal TV had the strongest association with Depersonalization ($r = 0.24$, $p < 0.01$). Emotional Exhaustion significantly associated with five of the six types of TVs, with Social TV having the strongest association ($r = 0.16$, $p < 0.01$), and Cyber TV having the weakest association ($r = 0.05$, p = ns.). Teachers' Sense of Accomplishment had the weakest associations with all six types of TVs, with the correlation coefficients ranging from -.08 with Cyber TV ($p < 0.05$) to – .04 with Physical TV (p = ns.) and Sexual Harassment (p = ns.). Also, strong correlations across subscales were found, with coefficients ranging from .71 to .96.

Discussion

The impact of teacher victimization (TV) on teachers' well-being has gained worldwide attention (Espelage et al., 2013; Wang, 2014); however, there are limited psychometrically-sound comprehensive measures, especially outside of the U.S. Thus, we developed the MTVS—the first multidimensional, psychometrically-sound measure to assess TV in China. Results indicated that the MTVS is best represented by a second-order factor structure, including six subscales and one general factor. This is consistent with the APA Task Force on Teacher-

Directed Violence (McMahon et al., 2014), which found six related forms of TV. The second-order model fit equally well across male and female teachers, as well as across middle and high school teachers.

Latent mean differences indicated that male teachers reported significantly higher levels of total TV, as well as social, verbal, and cyber TV and personal property offenses. Similar levels of physical and sexual TV were reported by male and female teachers. Previous research has found potential group differences, with male teachers reporting higher levels of physical victimization and female teachers reporting higher levels of nonphysical types (Berg & Cornell, 2016), possibly because male teachers may intervene with aggressive students, which may lead to increased physical TV for males (Martinez et al., 2016; McMahon et al., 2015). However, the majority of research examining gender differences has been conducted in the United States (e.g., Berg & Cornell, 2016; McMahon et al., 2015; Martinez et al., 2016) utilizing dichotomous measures; thus, additional research is needed. Similar levels of TV were reported across middle and high school teachers. This is inconsistent with previous research; however, many of these studies do not report statistical differences among grade levels (e.g., Chen & Astor, 2009; Kauppi & Pörhöla, 2012; Lyon & Douglas, 1999).

The significant and positive association between TV and teacher burnout found in the present study supported the concurrent validity of MTVS; it was also consistent with the findings in previous studies (Dzuka & Dalbert, 2007; Galand, Lecocg, & Philippot, 2007; Ozkilic & Kartal, 2012; Wilson et al., 2011). Interestingly, all TV types exhibited a more robust relation with depersonalization (i.e., distancing oneself from others) than with emotional exhaustion and accomplishment. Previous research has indicated that emotional exhaustion may be the first symptom of burnout; depersonalization may emerge to cope with emotional exhaustion (Boles,

Dean, Ricks, Short, & Wang, 2000). In the present study, teachers were asked to respond to the burnout survey based on their experience in the past school year. Because this information was not collected right after the specific victimization incident, it is possible that teachers may have passed the emotional exhaustion stage and instead are reflecting on coping strategies they used. Also, most teachers reported low frequency of TVs, thus the stress may not have accumulated to the extent that they are questioning their ability to teach.

Cyber victimization was not significantly correlated with emotional exhaustion; physical victimization and sexual harassment were not associated with accomplishment. Teachers may not feel that physically aggressive or sexual behaviors directed towards them are related to their career accomplishments, whereas verbal aggression (e.g., insults) may more directly relate to the teacher's sense of accomplishment. It may be more stressful to experience victimization directly during school (i.e., in front of students and colleagues), which may be why cyber victimization was not related to emotional exhaustion. Although these correlations are significant, they are indicative of small-to-moderate associations between TV and burnout, which may be due to the small percentage of teachers that reported TV. Results also indicated evidence of test-retest reliability for the MTVS, particularly for the total victimization scale. Reliability coefficients for some of the subscales were lower (e.g., physical victimization); this is likely due to the frequency of reported behaviors and the smaller sample utilized for these analyses ($n = 61$).

Limitations, Future Research, and Implications

Although the sample was large, caution is warranted in generalizing these results outside of a Chinese sample. Larger and more diverse samples are needed to validate the MTVS among other populations, countries, and demographic groups in China. Future studies should employ random sampling procedures. Further research is also needed to examine other psychometric

properties of the MTVS (e.g., convergent and predictive validity). Utilizing larger samples of teachers that report higher frequency of victimization may also be important. Although the purpose of the study was to develop and validate a teacher report measure, a multi-informant approach (e.g., student or administrator reports, behavior observations) may provide a more accurate and comprehensive picture of TV. The second-order model was chosen for the current study (over the four-factor model) because it was more parsimonious and aligned with theory; however, more evidence may be needed to support this model.

The MTVS can be utilized to comprehensively assess subtypes of TV, as well as general TV. This is the first study to validate a second-order model for a TV measure. Not only is this measure practical for research, it is also a useful tool for schools to assess prevalence and types of TV. Recent research has highlighted the importance of administrative support in TV prevention (McMahon, Reaves, McConnell, Peist, & Ruiz, 2017). The MTVS could be utilized to systematically assess TV school-wide. Utilizing systematic data collection could inform school-wide behavioral policies and trainings for teachers on fostering teacher-student relationships, discipline practices, and how to appropriately respond to aggressive behaviors.

Conclusion

The current study found evidence for the second-order model of the MTVS and multi-group invariance testing supported this factor structure across gender and grade levels. Male and female teachers reported different levels across the different types of victimization; however, middle and high school teachers reported similar levels of TV. Evidence of convergent validity was found with teacher burnout. These results provide preliminary psychometric support for the MTVS among a Chinese sample.

References

Aiken, L. S., Stein, J. A., & Bentler, P. M. (1994). Structural equation analyses of clinical subpopulation differences and comparative treatment outcomes: characterizing the daily lives of drug addicts. *Journal of Consulting and Clinical Psychology, 62*, 488-499. doi: 10.1037/0022-006X.62.3.488

Asparouhov, T., & Muthén, B. (Producer). (2010). Computing the strictly positive Satorra - Bentler chi-square test in Mplus. *Mplus Web Notes: No. 12* Retrieved from http://www.statmodel.com/examples/webnotes/webnote12.pdf

Berg, J. & Cornell, D. (2016). Authoritative school climate, aggression toward teachers, and teacher distress in middle school. *School Psychology Quarterly*, *31*, 122-127. doi:10.1037/spq0000132

Boles, J. S., Dean, D. H., Ricks, J. M., Short, J. C., & Wang, G. (2000). The dimensionality of the Maslach Burnout Inventory across small business owners and educators. *Journal of Vocational Behavior, 56*, 12-34. doi: 10.1006/jvbe.1999.1689

Bounds, C., & Jenkins, L. (2016). Teacher directed violence in relation to social support and stress. *Contemporary School Psychology, 20,* 1-9. doi:10.1007/s40688-016-0091-0

Byrne, B. M. (2006). *Structural equation modeling with EQS: Basic concepts, applications, and programming* (2nd ed.). Mahwah, NJ: Erlbaum.

Chen, J. K., & Astor, R. A. (2009). The perpetration of school violence in Taiwan: An analysis of gender, grade level and school type. *School Psychology International, 30*, 568–584. doi: 10.1177/0143034309107076

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9*, 233-255. doi:10.1207/s15328007sem0902_5

Dong, B. (2010). *A survey of middle school student of teacher bullying.* (Unpublished master's thesis). Central China Normal University, China.

Dzuka, J., & Dalbert, C. (2007). Student violence against teachers: Teachers' well-being and the belief in a just world. *European Psychologist, 12,* 253–260. doi:10.1027/1016-9040.12.4.253

Espelage, D., Anderman, E.M., Brown, V.E., Jones, A., Lane, K.L., McMahon, S.D., … Reynolds, C.R. (2013). Understanding and preventing violence directed against teachers: Recommendations for a national research, practice, and policy agenda. *American Psychologist, 68*, 75-87. doi: 10.1037/a0031307

Galand, B., Lecocg, C., & Philippot, P. (2007). School violence and teacher professional engagement. *British Journal of Educational Psychology, 77,* 465-477. doi:10.1348/000709906X114571

Guthrie, D. (2012). *China and globalization: The social, economic and political transformation of Chinese society.* New York: Routledge.

Hu, L. T., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods, 3*, 424-453. doi: 10.1037/1082-989X.3.4.424

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

Ingersoll, R., (2001). Teacher turnover and teacher shortages: An organizational analysis. *American Educational Research Journal, 38*, 499-534. doi:10.3102/00028312038003499

Kauppi, T., & Pörhölä, M. (2012). School teachers bullied by their students: Teachers' attributions and how they share their experiences. *Teaching and Teacher Education, 28*, 1059-1068. doi: 10.1016/j.tate.2012.05.009

Kowalski, R. M., Giumetti, G. W., Schroeder, A. N., & Lattanner, M. R. (2014). Bullying in the digital age: A critical review and meta-analysis of cyberbullying research among youth. *Psychological Bulletin, 140*, 1073-1137. doi:10.1037/a0035618

Lyon, D. R., & Douglas (1999). *Violence against British Columbia teachers: report of the Simon Fraser University/British Columbia Teachers' Federation violence against teachers survey*. Burnaby, B.C.: Mental Health, Law & Policy Institute, Simon Fraser University.

Marsh, H. W., Nagengast, B., Morin, A. J. S., Parada, R. H., Craven, R. G., & Hamilton, L. R. (2011). Construct validity of the multidimensional structure of bullying and victimization: An application of exploratory structural equation modeling. *Journal of Educational Psychology, 103*, 701-732. doi:10.1037/a0024122

Martinez, A., McMahon, S. D., Espelage, D., Anderman, E. M., Reddy, L.A., Sanchez, B. (2015). Teachers' experiences with multiple victimization: Identifying demographic, cognitive, and contextual correlates. *Journal of School Violence*, *15,* 387-405. doi:10.1080/15388220.2015.1056879

McCormick, J., & Barnett, K. (2011). Teachers' attributions for stress and their relationships with burnout. *International Journal of Educational Management, 25*, 278-293. doi: 10.1108/09513541111120114

McMahon, S.D., Martinez, A., Espelage, D., Rose, C., Reddy, L.A., Lane, K., … Brown, V. (2014). Violence directed against teachers: Results from a national survey. *Psychology in the Schools, 51,* 753-766. doi: 10.1002/pits.21777

Mooij, T. (2011). Secondary school teachers' personal and school characteristics, experience of violence and perceived violence motives. *Teachers and Teaching: Theory and Practice*, *17,* 227–253. doi: 10.1080/13540602.2011.539803.

Moon, B., Morash, M., Jang, J. O., & Jeong, S. (2015). Violence against teachers in South Korea: Negative consequences and factors leading to emotional distress. *Violence and Victims, 30*, 279-292. doi:10.1891/0886-6708.30.2.279

Muthén, B. O., & Satorra, A. (1995). Complex sample data in structural equation modeling. *Sociological Methodology, 25,* 267–316. doi:10.2307/271070

Mynard, H., & Joseph, S. (2000). Development of the multidimensional peer-victimization scale. *Aggressive Behavior, 26*, 169-178. doi: 10.1002/(SICI)1098-2337(2000)26:2<169::AID-AB3>3.0.CO;2-A

Ozkilic, R., & Kartal, H. (2012). Teachers bullied by their students: How their classes influenced after being bullied*? Procedia-Social and Behavioral Sciences, 46*, 3435-3439. doi: 10.1016/j.sbspro.2012.06.080

Reddy, L. A., Espelage, D. L., Anderman, E. M., Kanrich, J., & McMahon, S. D. (2018). Addressing Violence against Educators through measurement and Research. Unpublished Manuscript.

Thompson, B. (2004). *Exploratory and confirmatory factor analysis: Understanding concepts and applications*. Washington, DC: American Psychological Association.

Türküm, A. S. (2011). Social supports preferred by the teachers when facing school violence. *Children and Youth Services Review*, *33*, 644-650. doi:10.1016/j.childyouth.2010.11.005

Wang, C. (2014, December 18*).* Fear of violence from students prompts Yunnan teachers to go on strike. *The Paper*. Retrieved from www.thepaper.cn/newsDetail_forward_1286304

Wilson, C. M., Douglas, K. S., & Lyon, D. R. (2011). Violence against teachers: Prevalence and consequences. *Journal of Interpersonal Violence, 26*, 2353-2371. doi:10.1177/0886260510383027

Table 1

Observed Means and Differences in Latent Means of First-order and Second-order Factors

	Latent Mean Differences$_{Group1 - Group2}$				Observed means	
	Estimates	z	*d*	*d* (95% CIs)	*M (SD)*$_{Group 1}$	*M (SD)* $_{Group 2}$
Gender Differences (Group 1 = Male, Group 2 = Female)						
1. Total TV	-.16*	-2.39	- 0.13	[- 0.40, - 0.02]	1.11 (0.45)	1.05 (0.29)
2. Physical TV	-. 0.13	-1.62	- 0.11	[- 0.35, 0.00]	1.09 (0.48)	1.04 (0.31)
3. Social TV	- 0.19*	-2.26	- 0.15	[- 0.36,- 0.04]	1.17 (0.51)	1.10 (0.36)
4. Verbal TV	- 0.19*	-2.19	- 0.15	[- 0.41, - 0.03]	1.17 (0.51)	1.10 (0.35)
5. Cyber TV	- 0.20*	-2.14	- 0.15	[- 0.41, - 0.03]	1.11 (0.53)	1.04 (0.30)
6. Sexual TV	-0.11	-1.44	- 0.09	[- 0.34, 0.01]	1.12 (0.50)	1.06 (0.33)
7. Personal Property Offenses	- 0.16*	-2.01	- 0.14	[- 0.38, - 0.02]	1.12 (0.48)	1.06 (0.33)
Grade-level Differences (Group 1 = Middle School, Group 2 = High School)						
1. Total TV	0.02	0.72	0.04	[- 0.12, 0.13]	1.08 (0.36)	1.09 (0.46)
2. Physical TV	0.02	0.66	0.04	[- 0.13, 0.12]	1.06 (0.40)	1.07 (0.47)
3. Social TV	0.03	0.98	0.06	[- 0.09, 0.15]	1.12 (0.42)	1.15 (0.52)
4. Verbal TV	0.03	1.03	0.06	[- 0.09, 0.15]	1.13 (0.42)	1.15 (0.50)
5. Cyber TV	0.02	0.72	0.04	[- 0.12, 0.13]	1.07 (0.40)	1.08 (0.47)
6. Sexual TV	0.02	0.84	0.05	[- 0.12, 0.13]	1.08 (0.40)	1.13 (0.51)
7. Personal Property Offenses	0.01	0.59	0.04	[- 0.14, 0.12]	1.09 (0.40)	1.10 (0.49)

*$p < .001$.

Table 2

Correlations between Teacher Victimization (TV) and Teacher Burnout

	Total Burnout	Emotional Exhaustion	Accomplishment	Depersonalization
1. Total TV	.19**	.09**	- .06*	.21**
2. Physical TV	.15**	.06*	- .04	.19**
2. Social TV	.27**	.16**	- .07**	.27**
3. Verbal TV	.23**	.13**	- .07**	.24**
4. Cyber TV	.14**	.05	- .08**	.15**
5. Sexual Harassment	.13**	.06*	- .04	.16**
6. Personal Property Offenses	.19**	.09**	- .05*	.21**

Note. TV = Teacher Victimization. **$p < .01$. *$p < .05$.

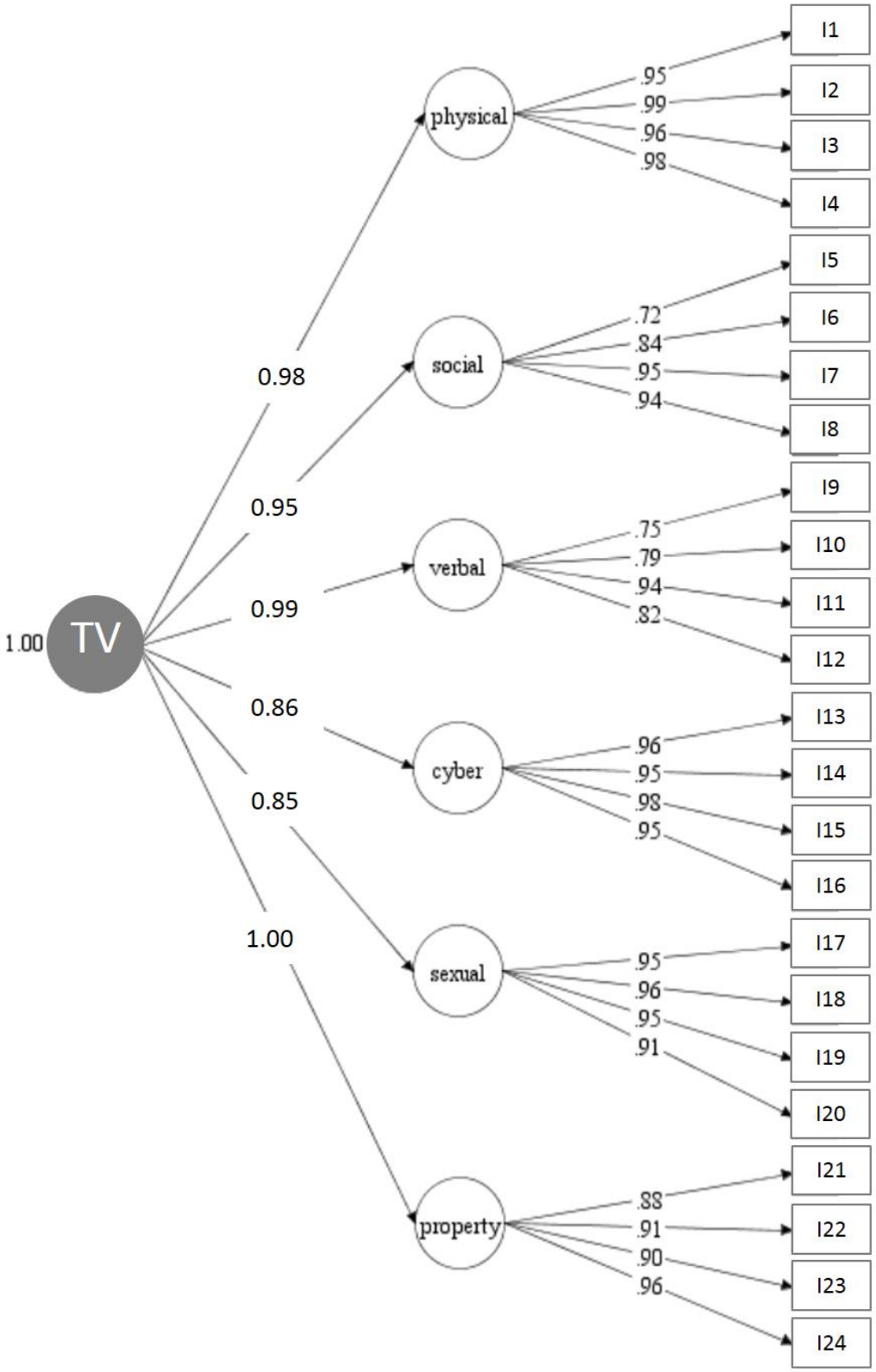

Figure 1. Standardized factor loadings for the second-order model in the full sample