



Psychosocial Intervention

<https://journals.copmadrid.org/pi>



Assessment of a Multidimensional School Collective Efficacy Scale to Prevent Student Bullying: Examining Dimensionality and Measurement Invariance

José N. Peraza-Balderrama^a, Angel A. Valdés-Cuervo^a, Belén Martínez-Ferrer^b, Ana C. Reyes-Rodríguez^c, and Lizeth G. Parra-Pérez^d

^aTechnological Institute of Sonora, Mexico; ^bPablo de Olavide University, Sevilla, Spain; ^cIrapuato Institute, Guanajuato, Mexico; ^dColorado State University, Fort Collins, USA

ARTICLE INFO

Received 22 September 2020
Accepted 22 January 2021
Available online 5 March 2021

Keywords:

School collective efficacy
Bullying
Internal structure
Measurement invariance
External validity

ABSTRACT

The construct of a school collective efficacy to prevent bullying has attracted attention as a way to increase a positive, school-wide climate. The current study tested the fit of several first-order models of school collective efficacy to prevent (uni-dimensional, two-dimensional, and three-dimensional) bullying using a sample of 579 male ($M_{age} = 14.31$, $SD = 1.78$ years old) and 589 female ($M_{age} = 14.56$, $SD = 1.83$ years old) Mexican adolescents. The models were validated by the extent to which the model was invariant by gender and by educational level (secondary vs. high school). Moreover, the discriminant and concurrent validity of model dimensions were examined through their relationships with other constructs. The results suggest that school collective efficacy is a three-dimensional construct, with supporting evidence for cohesion, students' social control, and teachers' social control dimensions. Measurement invariance was found in this three-dimensional measurement model by gender and educational level. The latent means difference analysis showed some differences by gender and educational level on factors of school collective efficacy. Finally, results support our hypotheses related to discriminant and concurrent validity in relation to external variables. Overall, findings indicate this three-dimensional model is useful to measure adolescents' perceptions of school collective efficacy.

Una evaluación de la *Multidimensional School Collective Efficacy Scale* para prevenir el acoso en escolares: el examen de la dimensionalidad y la invarianza de medida

RESUMEN

El constructo eficacia escolar colectiva para la prevención del acoso escolar ha recibido una creciente atención como una forma de mejorar el clima escolar. En el presente estudio se evaluó el ajuste de varios modelos de primer orden para medir la autoeficacia escolar colectiva para la prevención del acoso escolar (unidimensional, bidimensional y tridimensional), utilizando una muestra de 579 chicos ($M_{edad} = 14.31$, $DS = 1.78$ años) y 589 chicas ($M_{edad} = 14.56$, $DS = 1.83$ años) mexicanos. Para validar el modelo en distintos grupos se analizó la invarianza de medida por sexo y nivel educativo (secundaria vs. bachillerato). Además, se examinó la validez discriminante y concurrente analizando la relación de la eficacia escolar colectiva con otros constructos. Los resultados indican que la eficacia escolar colectiva es un constructo tridimensional conformado por las dimensiones de cohesión y confianza, control social de los estudiantes y control social de los docentes. La evidencia constata la invarianza de medida del modelo por sexo y nivel educativo. El análisis de las diferencias de medias latentes muestra diferencias por sexo y nivel educativo en las dimensiones de la eficacia escolar colectiva. Finalmente, los resultados confirman las hipótesis relativas a las relaciones de la eficacia escolar colectiva con otros constructos. Los hallazgos sugieren que el modelo tridimensional es útil para medir la percepción adolescente de la eficacia escolar colectiva.

Palabras clave:

Eficacia escolar colectiva
Acoso escolar
Estructura interna
Invarianza de medida
Validez externa

Bullying is an aggressive, goal-directed behavior that occurs when someone intentionally inflicts physical, emotional, or social injuries upon another. This behavior is carried out repeatedly and over time

in interpersonal relationships, with unequal power between victim and aggressor (Olweus, 1993; Volk et al., 2014). Bullying is associated with short-term and long-term consequences, such as lower academic

Cite this article as: Peraza-Balderrama, J. N., Valdés-Cuervo, A. A., Martínez-Ferrer, B., Reyes-Rodríguez, A. C., & Parra-Pérez, L. G. (2021). Assessment of a multidimensional school collective efficacy scale to prevent student bullying: Examining dimensionality and measurement invariance. *Psychosocial Intervention*. Ahead of print. <https://doi.org/10.5093/pi2021a2>

Correspondence: angel.valdes@itson.edu.mx (A. A. Valdés-Cuervo).

ISSN:1132-0559/© 2021 Colegio Oficial de la Psicología de Madrid. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

outcomes, psychological disorders, and behavioral problems (du Plessis et al., 2019; Garaigordobil & Machimbarrena, 2019; Nielsen et al., 2015; Sigurdson et al., 2015; Sourander et al., 2016). It is, in fact, a public health issue faced worldwide (Modecki et al., 2014; United Nations Educational, Scientific and Cultural Organization [UNESCO, 2017]; Zych et al., 2017). Some researchers have argued bullying is a pervasive behavior that occurs in social contexts, regardless of whether it is promoted, supported, or rejected (Espegale et al., 2013; Martínez-Ferrer et al., 2011; Swearer & Hymel, 2015). However, other scholars (de Oliveira et al., 2015; Hymel et al., 2015; Roland & Idsøe, 2001; Valdés-Cuervo et al., 2018) have found that school-level variables are responsible for differences in bullying rates between schools. Specifically, research has highlighted school collective efficacy as a whole-level school factor is associated with lower rates of peer aggression (Barchia & Bussey, 2011; Olsson et al., 2017; Sapouna, 2010; Thornberg et al., 2019; Williams & Guerra, 2011).

School Collective Efficacy

The collective efficacy theory developed from the study of violence in neighborhoods (Sampson & Raudenbush, 1999; Sampson et al., 1997). It is rooted in social capital (Bourdieu, 1984; Coleman, 1990) and social cognitive theories (Bandura, 1986, 1989). It aimed to examine the mechanism through which the social context might affect individual behavior. The collective efficacy theory acknowledges that violence is associated with the social resources that the community holds, as well as the beliefs shared by its members (Häuberer et al., 2011; Sampson et al., 1997; Wilcox et al., 2018). Hence, social resources and shared beliefs are critical to collective agency (Bandura, 2000, 2001). Therefore, violence prevention in schools relies on the whole-school context, rather than only on the students themselves.

In this regard, some scholars (Coleman, 1990; Sampson, 2006) assert that friendships and fellowships are not enough to avoid violent episodes inside schools. They argue that students must feel a strong tie and belong to 'the group' in order to take co-responsibility for its success and to protect its members. According to the collective efficacy theory, shared feelings of cohesion and trust, along with a willingness of members to exert informal social control, lead to the values of the community, but also sets the goals in terms of violence prevention (Sampson et al., 1997; Wilcox et al., 2018). Cohesion and trust refer to the level of confidence and support among members of a group, who act accordingly to accomplish common goals (Sampson et al., 1999). Informal social control refers to the willingness of community members to intervene for the common good (Sampson et al., 1997). It involves informal practices and arrangements that contribute to the institutional and community order (Black, 1984; Taylor, 2001). Informal social control is particularly effective because it reduces coercion (Hirschi, 2001; Janowitz, 1975; Wiatrowski et al., 1981), that is, results are more effective when leading individuals of a community embrace collective value goals, rather than forcing them to adjust themselves to a social reality.

In bullying research, school collective efficacy is conceptualized as school members' beliefs about the cohesion and trust between students and adults at school, as well as the willingness of school members to support victims of bullying (Hymel et al., 2015; Williams & Guerra, 2011). School cohesion and informal social control have influence on both adults and peers to provide support when bullying occurs (Capone et al., 2018; Kurnianingshi et al., 2012; Payne & Gottfredson, 2004; Thornberg & Elvstrand, 2012; Trach et al., 2013; Wach et al., 2018; Williams & Guerra, 2011).

Measures of Students' Perception of School Collective Efficacy

Although researchers have developed scales for measuring school collective efficacy as a means to predict bullying episodes, most

are considered either unidimensional (e.g., Barchia & Bussey, 2011) or multidimensional (physical, relational, and verbal aggression) indicators of social informal control (e.g., Wänström et al., 2017). That is, none of these scales have included indicators of cohesion and trust. While current scales allow researchers to assess overall information on school collective efficacy, this type of measurement does not provide insights into the theoretical multidimensionality of the construct (Sampson & Raudenbush, 1999; Sampson et al., 1997). Thus, if school collective efficacy is a multidimensional construct that comprises cohesion, trust, and social informal control, then, the existing instrument is limited to the exploration of differences between school efficacy dimensions. More recently, scholars (e.g., Sapouna, 2010; Williams & Guerra, 2011) have developed other scales to address the lack of a theoretical multidimensional model; these scales, however, have presented limited evidence about their psychometric properties.

Specifically, Sapouna (2010) developed a two-dimensional instrument to measure social cohesion and trust, and informal social control with Greek students ranging in age from 11 to 14 years. However, the author did not report the psychometric properties of the scale. In the case of Williams and Guerra (2011), their three-dimensional instrument had several important issues as well. Scholars tested American pre-adolescents and adolescents by running an exploratory factor analysis with principal components extraction and oblique rotation, which supported a three-dimensional (cohesion and trust, students' informal social control, and teachers' informal social control) solution (67% explained variance). Nonetheless, their results should be taken with caution since the principal component analysis (PCA) method has an important weakness: PCA assumes that all indicator variance is common. The assumption is rigorous and does not allow for specific variance; that is, it assumes that scores are perfectly reliable (Kline, 2013; Park et al., 2002). Also, PCA only analyzes observed variables; therefore, it should not be used if scholars seek to obtain parameters reflecting latent constructs (Gaskin & Happell, 2014; Widaman, 1993). Finally, Williams and Guerra (2011) do not report the correlation between the factors, which makes it difficult to analyze the nature of the relationship between factors and the scoring of the scale.

After conducting a literature review, more important gaps became apparent. First, there is a dearth of studies comparing the adjustment of different measurement models of the construct (school collective efficacy). Second, no study known by the authors has analyzed the measurement invariance of the proposed model with relevant variables, such as gender and stage of development (e.g., childhood, pre-adolescence, adolescence, adulthood). Third, research about the external construct validity of measures of school collective efficacy, remains scarce. And fourth, to our knowledge, no previous studies have analyzed the psychometric properties of a scale to measure collective school efficacy in Latin American adolescents, specifically.

Considering all of the above, in an effort to address these gaps in the current body of literature, we sought to develop an instrument capable of measuring the three dimensions of school collective efficacy (cohesion and trust, students' informal social control, and teachers' informal social control). We drew on the work of Sampson and Raudenbush (1999), and Sampson et al. (1997), to develop a theoretically grounded and psychometrically robust scale, capable of measuring the different indicators of school collective efficacy in Mexican adolescents.

Measurement Invariance

Prior studies (Barchia & Bussey, 2011; Marées & Petermann, 2010) suggest student perception of school collective efficacy differs by gender and educational level. Overall, girls report higher perceptions of school collective efficacy in bullying prevention

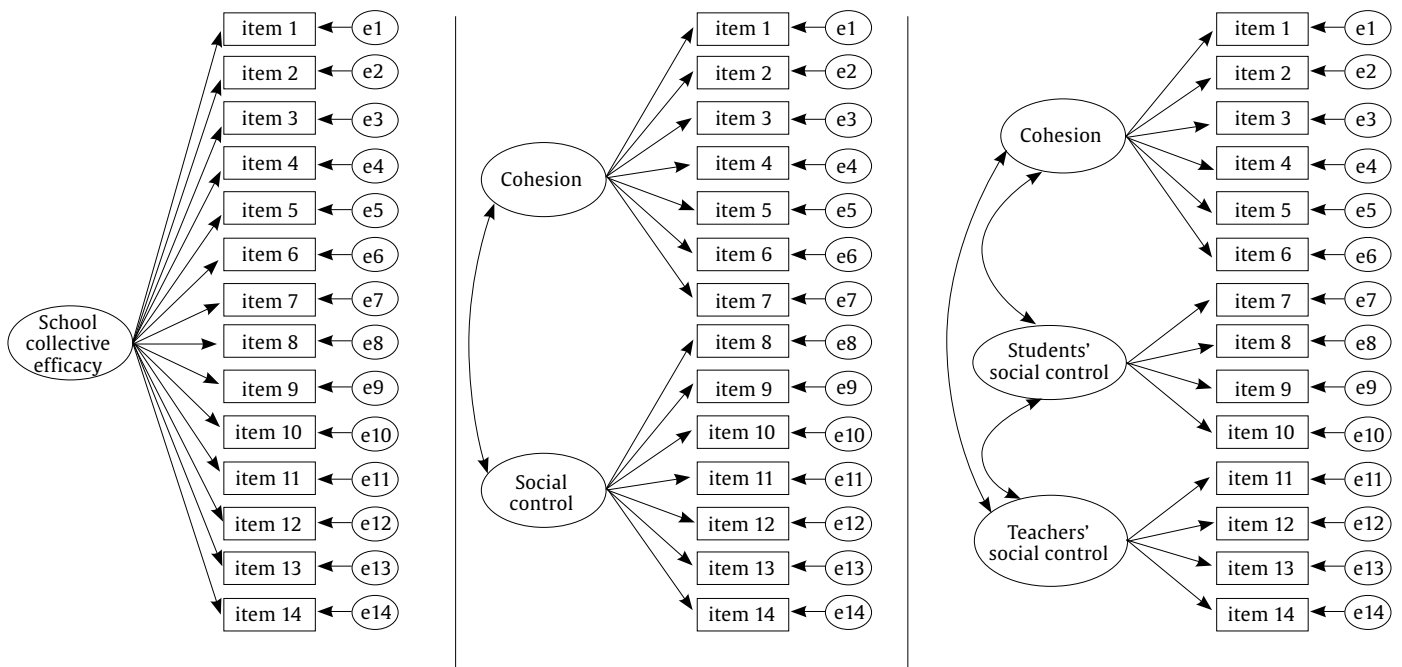


Figure 1. Factor Model of School Collective Efficacy Depicting One-Dimensional, Two-dimensional, and Three-dimensional Structures.

than boys throughout schooling (Barchia & Bussey, 2011; Cho, 2017; Lee & Jo, 2017). Nonetheless, research on school collective efficacy has produced limited evidence on age differences, with a single study found by the authors reporting that secondary students perceive a higher school collective efficacy than high school students (Barchia & Bussey, 2011). These studies, however, did not consider the effects of measurement invariance when making group comparisons, so it is still unknown to what extent school collective efficacy is interpreted in the same way across gender and educational level (Brown, 2015; Millsap & Olivera-Aguilar, 2012). Thus, if school collective efficacy is different across gender and educational level, and these variations are not considered, it is inadequate to compare school collective efficacy across these groups (Millsap & Olivera-Aguilar, 2012; Putnick & Bornstein, 2016). Therefore, it is essential to examine whether the measurement of school collective efficacy is similar by gender and stage of adolescence. Without confirming measurement invariance, group comparisons are prone to be misrepresented. Measurement invariance allows a comparison of school collective efficacy among these groups or common scores that can be used to predict school bullying or other variables (Byrne, 2016; Van de Schoot et al., 2015).

Concurrent Validity

Scholars suggest that the influence of school collective efficacy on bullying prevention is explained by its relationship with whole-school variables (Hymel et al., 2015; Williams & Guerra, 2011). Although research about benefits or consequences of school collective efficacy on school environment is scarce, related studies (e.g., Barchia & Bussey, 2011; Capone et al., 2018; Thornberg et al., 2019; Thornberg et al., 2020) show that collective efficacy is associated to indicators of quality of school environment. Thus, to assess the external validity of the scale, we analyzed relationships of school collective efficacy with whole-school level variables, such as school collective moral disengagement (Barchia & Bussey, 2011; Gini et al., 2015; Pozzoli et al., 2012; Thornberg et al., 2019), sense of school community (Capone et al., 2018; Hymel et al., 2015; Prati et al., 2017), and school fairness

(Konishi et al., 2017; Lenzi et al., 2014; Vieno et al., 2013).

The literature defines collective moral disengagement as the perception of individuals about the inclination of school members to use cognitive mechanisms to justify immoral actions (Allison & Bussey, 2017; Bandura, 2002). In this regard, some studies (Barchia & Bussey, 2011; Gini et al., 2015; Pozzoli et al., 2012; Thornberg et al., 2019) report a negative association between school collective efficacy and collective moral disengagement in bullying studies. These results suggest that school collective efficacy to prevent bullying leads school members' moral responsibility in immoral behaviors. In the case of sense of school community, some scholars (Albanesi et al., 2007; Prati et al., 2017) argue that school collective efficacy spurs sense of school community in bullying prevention. Albanesi et al. (2017), explain that the sense of community involves students' perceptions of emotional connection with peers and school staff, and opportunities to participate in school life. In fact, it is argued that personal needs are satisfied through such school membership (Albanesi et al., 2007; Prati et al., 2017). Hence, school collective efficacy promotes school fairness, as it prompts students to perceive school rules and discipline practices as just and fair (Peguero et al., 2017; Peterson & Seligman, 2004).

The Present Study

Prior studies have reported a significant effect of school collective efficacy on reducing peer aggression and bullying (Barchia & Bussey, 2011; Olsson et al., 2017; Sapouna, 2010; Thornberg et al., 2019; Williams & Guerra, 2011). However, there are important gaps in the measurement of student perceptions of collective school efficacy. Under this context, this study sought to: (1) analyze the internal structure of validity by examining the fit of different measurement models of the construct: only one factor, two first-order factors (cohesion and trust and social control), and three first-order factors (cohesion and trust, students' social control, and teachers' social control) (see Figure 1; see Appendix); (2) test reliability by running average variance extracted (AVE) and internal consistency of scores; (3) explore measurement invariance by gender and educational level (secondary vs. high school); (4) compare latent variable

mean differences across groups if measurement invariance was present; and (5) analyze (discriminant and concurrent) external validity, by analyzing its relationships with collective school moral disengagement, sense of school community, and school fairness.

To accomplish these goals, five hypotheses were explored. Hypothesis 1 (internal structure): the three first-order factor structure has a better adjustment to the data than the one or two first-order factor models. Hypothesis 2 (measurement invariance): the scale is invariant by gender and educational level. Hypothesis 3 (means comparison): girls hold a greater perception of school collective efficacy than boys. Given the scarcity of studies on education level group differences, we make no specific hypotheses regarding these differences. Hypothesis 4 (discriminant validity): each subscale discriminates from the other subscales of the school collective efficacy scale. And, hypothesis 5 (concurrent validity): school collective efficacy has positive relations with sense of school community and school fairness, and a negative one with school collective moral disengagement.

See complete item question in [Appendix](#).

Method

Participants

Participants came from urban public secondary ($n = 83$) and high schools ($n = 30$) across different cities in the state of Sonora, Mexico (Hermosillo, Obregón, and Guaymas). The sample of secondary students included 584 participants (6 to 8 students per school), 300 males and 284 females, aged from 11 to 14 years old ($M_{\text{age}} = 12.67$ years old, $SD = .77$ years old); 204 (35%) of them were enrolled in seventh grade, 187 (32%) in eighth grade, and 193 (33%) were in ninth grade. The sample from high school (17 to 20 students per school) included 279 (48.3%) males and 305 (51.7%) females who ranged from 14 to 17 years old ($M_{\text{age}} = 15.95$ years old, $SD = .78$ years); 187 (32%) attended tenth grade, 210 eleventh grade, and 187 (32%) twelfth grade. Like most public urban schools in Mexico, these schools included students from low and middle socio-economic status (National Institute for the Evaluation of Education [[INEE, 2018](#)]).

Measures

School collective efficacy to bullying prevention. The work conducted by previous authors ([Barchia & Bussey, 2011](#); [Sapouna, 2010](#); [Williams & Guerra, 2011](#)) was used to develop a self-reported scale. After building 30 items aimed to measure school collective efficacy, we sought assistance from five knowledgeable researchers in the field of bullying and two school counselors as well. Specifically, they were asked to assess the relevance of each item. The experts rated the relevance of each item on a 4-point Likert-type scale (e.g., "Is this item relevant to measure school collective efficacy?"; 1 = *not relevant*, 2 = *unable to assess relevance without item major revision*, 3 = *relevant*, and 4 = *very relevant*). Through this process, 14 items were most appropriate and included in the School Collective Efficacy Scale (SCEE), as they had a context validity index (CVI) greater than .78 ([Almanasreh et al., 2019](#); [Lynn, 1986](#)).

Included items are indicators of cohesion (e.g., "I believe my school has achieved agreement from students to end bullying"); informal students' social control (e.g., "Students at my school stop bullying when a student is victim of teasing from a stronger student"); and informal teachers' social control (e.g., "Teachers at my school stop bullying when a student is victim of rumors and lies"). Research participants responded using a four-point Likert scale from (from 0 = *strongly disagree*, to 4 = *strongly agree*). Items were coded so that higher scores indicate higher school collective efficacy to prevent bullying.

School collective moral disengagement. The Classroom Collective Moral Disengagement Scale (CMD; [Kollerová et al., 2018](#)) was used. The back-translation method was used for the appropriateness of the scale to Mexican adolescents. This scale is a uni-dimensional measurement model with 10 items (e.g., "In this school, how many students believe that it is all right to fight to protect friends," $\alpha = .81$, $\omega = .82$, AVE = .62). Students were asked about "In your school, how many students think that" in a 5-point scale (0 = *none 0%*, to 4 = *everyone 100%*). The CFA indicated a good model fit to the data ($SB\chi^2 = 48.51$, $df = 31$, $p = .024$; SRMR = .03; AGFI = .98; CFI = .99; TLI = .99; RMSEA = .02, CI 90 [.01, .04]).

Sense of school community. The Scale of the Sense of Community in the School ([Prati et al., 2017](#)) was used. The back-translation method was used for the translation of the scale from English to Spanish. The scale included 11 items grouped in three dimensions: (a) membership (3 items, e.g., "I'm proud to belong to this school," $\alpha = .78$, $\omega = .79$, AVE = .51), (b) emotional connection (3 items, e.g., "I feel that I can share experiences and interest with other students at my school," $\alpha = .79$, $\omega = .81$, AVE = .53), and (c) opportunities (5 items, e.g., "The students at my school are involved in organizing school events", $\alpha = .82$, $\omega = .82$, AVE = .57). The CFA supported a good model fit ($SB\chi^2 = 57.24$, $df = 31$, $p = .003$; SRMR = .04; AGFI = .97; CFI = .99; TLI = .98; RMSEA = .03, CI 90 [.01, .04]).

School fairness. The subscale of fairness adapted the International Personality Item Pool (IPIP; [du Plessis & Bruin, 2015](#)). The back-translation method was used for the translation of the scale from English to Spanish. Then, the scale was adjusted so that the student refers to the whole-school instead of personal fairness. The subscale comprises 9 items (e.g., "In this school, students admit when they are wrong," $\alpha = .86$, $\omega = .86$, AVE = .63). Likert scale responses with options from 0 = *totally disagree*, to 4 = *totally agree* were used. The CFA supported that the model fit the data ($SB\chi^2 = 50.28$, $df = 24$, $p = .002$; SRMR = .03; AGFI = .95; CFI = .97; TLI = .96; RMSEA = .05, CI 90 [.03, .06]).

Procedure

Researchers first gained permission from the Ethical Committee of the Technological Institute of Sonora (Number 2020_0014). An invitation to participate in the study was sent, through the education authorities, to the principals from different secondary and high schools from three cities of the state of Sonora, Mexico (Hermosillo, Obregon City, and Guaymas). Schools with directors that expressed their interest in participating were included in the study. Then, a consent letter was sent to potential research participants' parents to request their children participate in the study. Only 6% of parents refused to allow their children to participate. We explained the purpose of the study to research participants; we also underlined their participation would always be voluntary. They were also informed about the anonymity of their participation, which was ensured throughout the procedures. Data collection was carried out grouping students from each institution with sessions of about 45 minutes.

A back-translation method was used to ensure accuracy for the translation of the scales from English to Spanish. A bilingual Mexican researcher translated scales from English to Spanish. Then, a professional translator was asked to complete the Spanish translation back to English. Two other translators then compared that new translation with the original text. Finally, any differences between the two versions were reconciled in the translator focus group and elaborated into the final Spanish version.

Data Analysis

The percentage of missed data was 3%. Missing data was treated by the means of multiple imputation, which is available in SPSS 25. We examined the hypothesized factor structure of the measurement

of school collective efficacy and its invariance by gender and educational level (secondary vs. high school). The analyses were conducted using the robust maximum likelihood estimator (MLR) in Mplus 8. Mplus was chosen due to the sample having students nested within schools, and Mplus is able to account for the nested nature of the data by fitting a multilevel model with random-effects and produces adjusted standard errors in the model estimations.

Dimensionality analysis. To examine the factorial validity of the SCEE, several confirmatory factor analyses (CFA) were calculated. The models described above were analyzed and compared (uni-dimensional, two-dimensional, and three-dimensional factor model). In evaluating the goodness of fit of the hypothesized models with MLR estimator, we used the χ^2 statistic based on the Satorra-Bentler scaled statistic ($SB\chi^2$; Satorra & Bentler, 2010). The $SB\chi^2$ statistic incorporates a scaling correction for the χ^2 statistic (Hu et al., 1992). In addition, we reported the standardized root mean square error of approximation (SRMR $\leq .05$), comparative fit index (CFI $\geq .95$), Tucker-Lewis index (TLI $\geq .90$), and root mean square error of approximation (RMSEA $\leq .05$) with their confidence interval (Byrne, 2016; Sharma et al., 2005). In order to compare models, we used differences in $SB\chi^2$ ($\Delta SB\chi^2$) and Bayesian information criterion (BIC). We assumed that when differences in the $\Delta SB\chi^2$ result was significant, a model with less $SB\chi^2$ had better adjustment to the data (Byrne, 2016; Wang & Wang, 2012). Differences in BIC > 10 indicates differences in the model's fit to the data, a model with less BIC has strong adjustment (Byrne, 2016; Muthén & Muthén, 2017; Raftery, 1995).

Reliability analysis. Reliability of each scale was calculated using the McDonald coefficient (ω), composite reliability (CR), and average variance extracted (AVE). Values of ω and CR $\geq .70$ and AVE $\geq .50$ were considered indicators of an acceptable reliability of the scores (Bacon et al., 1995; Green, 2015; Hair et al., 2017).

Measurement invariance analysis. Nested factor models were fit in order to test measurement invariance, following the methods proposed in the literature (Millsap & Olivera-Aguilar, 2012; Putnick & Bornstein, 2016; Widaman & Resie, 1997). We tested configural invariance (constrained the number of factor and pattern of fixed and free factor loading to be the same across gender and educational level), metric invariance (constrained factor loadings to be equal across groups), and scalar invariance (constrained measurement intercept across groups). The nested models were compared by using the difference in the $\Delta SB\chi^2$, albeit with a correction to the value, as this difference is not distributed as a χ^2 statistic. Differences in $\Delta SB\chi^2$ greater than the critical value ($\Delta SB\chi^2$ with $p > .001$) suggests that constraints imposed are not equivalent across groups (Byrne, 2016; Muthén & Muthén, 2017). Nevertheless, the $\Delta SB\chi^2$ statistic is sensitive to larger samples sizes (Tomarken & Waller, 2005; Wang & Wang, 2012); thus, past researchers have recommended to use goodness-of-fit indexes, such as ΔCFI and $\Delta RMSEA$ (Byrne, 2016; Cheung & Rensvold, 2002; Sass & Smith, 2013). These guidelines were relied on for this study, with a difference in CFI (ΔCFI) less than .01, and differences in RMSEA ($\Delta RMSEA$) less than .015. In cases where methods disagreed, we relied more on differences in CFI and RMSEA due to the large sample size in this study.

Based on evidence of invariant factor loading and intercepts, we tested mean differences by gender and educational level (secondary vs. high school). For this, the factor means for the reference group (boys and secondary students) were set to zero, while we estimated groups' factor means freely. A z statistic was used to test differences between the latent means of the two groups (Aiken et al., 1994; Byrne, 2016).

Discriminant validity analysis. Discriminant validity requires evidence that the hypothesized differences between the constructs are empirically grounded (Farrell, 2010). Therefore, it requires that a latent variable can account for more variance in the observed variable associated with it than other constructs within a similar framework (Fornell & Larcker, 1981). For the purpose of this study, we used the

criterion that average variance extracted (AVE) in each SCE subscale is greater than the square of this correlation with the other subscales (Fornell & Larcker, 1981; Hair et al., 2010).

Concurrent validity analysis. Concurrent validity evidence refers to which scale scores are correlated in an expected manner with other constructs that are measured at the same time (Furr & Bacharach, 2014). To examine concurrent validity of the SCEE, correlations between the constructs involved in the study were calculated. Then, the correlation between school collective efficacy was calculated with similar (school fairness and school sense of community) and different constructs (school collective moral disengagement). Also, we evaluated the effect-size of these correlations based on the guidelines proposed by Funder and Ozer (2019). They suggested that an effect-size r of .10 is small for single events but potentially consequential, an effect-size r of .20 reveals a medium effect that is of some practical use in both the short and long term, and an effect size r of .30 suggests a large effect that is influential in the explanation in both short and long term.

Results

Descriptive Analysis and Multivariate Normality

Table 1 shows mean, standard deviation, range, skew, and kurtosis statistics. In most of the items (10 items), means are centered in the *neither agree nor disagree* middle category), while in the remaining (4 items) it is centered in the *disagree* category. These findings imply that respondents did not 'agree' with any of the item statements. Regarding the values of skewness and kurtosis, these suggest normal univariate distribution in all items. Finally, the Mardia coefficient was 4.98, which suggests multivariate normality.

Table 1. SCEE Items Descriptive Statistics

	<i>M</i>	<i>SD</i>	Min	Max	Skew	Kurtosis
Item 1	2.47	1.05	0	4	-0.51 (0.07)	0.12 (0.09)
Item 2	2.66	1.03	0	4	-0.72 (0.07)	0.23 (0.09)
Item 3	2.23	1.30	0	4	-0.21 (0.07)	-1.02 (0.09)
Item 4	2.65	0.95	0	4	-0.58 (0.07)	0.19 (0.09)
Item 5	2.66	1.23	0	4	-0.59 (0.07)	0.63 (0.09)
Item 6	2.46	1.19	0	4	-0.41 (0.07)	0.66 (0.09)
Item 7	1.78	1.22	0	4	-0.09 (0.07)	-0.94 (0.09)
Item 8	1.73	1.23	0	4	0.13 (0.07)	-0.98 (0.09)
Item 9	1.66	1.17	0	4	0.19 (0.07)	-0.75 (0.09)
Item 10	1.75	1.27	0	4	0.11 (0.07)	-1.04 (0.09)
Item 11	2.54	1.25	0	4	0.61 (0.07)	-0.59 (0.09)
Item 12	2.11	1.21	0	4	0.14 (0.07)	-0.91 (0.09)
Item 13	2.06	1.24	0	4	0.16 (0.07)	-0.91 (0.09)
Item 14	2.64	1.31	0	4	0.73 (0.07)	-0.62 (0.09)

Note. Min = minimum; Max = maximum.

Assessing Measurement Models

Confirmatory factor analysis of student responses to the 14 items verified that only the three-dimensional model fits best to the data. We compared the three-dimensional model (Model C) with the uni-dimensional (Model A), and the fit of Model A was worse to a statistically significant degree than Model C ($\Delta SB\chi^2 = 493.93$, $df = 3$, $p < .001$; $\Delta BIC = 261.44$). We further compared the three-dimensional model (Model C) and the two-dimensional model (Model B), the results show that Model B was also worse than Model C with statistical significance ($\Delta SB\chi^2 = 118.79$, $df = 2$, $p < .001$; $\Delta BIC = 155.8$). Therefore, based on theory and goodness of fit of the model, Model C was the preferred model, with the remaining analytical report based on this model (Table 2).

Table 2. Goodness-of-Fit Statistics of the Hypothesized One-Factor, Two-Factors, and Three-Factors Models ($N = 1,168$)

Model	$SB\chi^2$	df	p	SRMR	CFI	TLI	RMSEA	BIC	Comparison	$\Delta SB\chi^2$	Δdf	ΔBIC
A. One-factor	546.70	33	< .001	.14	.68	.65	.150	322.75	1 vs. 3	493.93	3	261.44
B. Two-factors	171.56	32	< .001	.07	.91	.90	.078	217.11	2 vs. 3	118.79	2	155.80
C. Three-factors	52.77	30	.006	.04	.99	.96	.003	61.31	-	-	-	-

Note. In Model A, all 14 items were loaded on one factor; in Model B, 6 items were loaded onto one factor (cohesion and trust), and the remaining 8 onto the other factor (informal social control); in Model C, 6 items were loaded in one factor (cohesion), 4 were loaded onto a second factor (informal students' social control), and the last 4 items were loaded onto a third factor (informal teachers' social control).

The estimated standardized factor loadings for the three-dimensional model is shown in Figure 2. The standardized factor loadings ranged from .63 to .83, and were statistically significant at the .05 level. The three-factor model provided the most acceptable fit to the data ($SB\chi^2 = 52.77$, $df = 30$, $p = .006$; $SRMR = .04$, $TLI = .96$; $CFI = .99$; $TLI = .98$; $RMSEA = .03$ IC 90 [.01, .04]). Furthermore, patterns in the bivariate correlations of the factors were found. Cohesion, students' social control, and teachers' social control were positively correlated. The correlations among the six factors were moderate to high (.28-.44). The reliability statistic for each latent variable is acceptable: cohesion (McDonald omega, $\omega = .81$, composite reliability $CR = .89$, and average variance extracted $AVE = .58$), students' social control ($\omega = .72$, $CR = .79$, $AVE = .50$), and teachers' social control ($\omega = .81$, $CR = .82$, $AVE = .54$). In sum, these results provide empirical evidence for measuring school collective efficacy with a three-dimensional model, rather than uni-dimensional or two-dimensional construct.

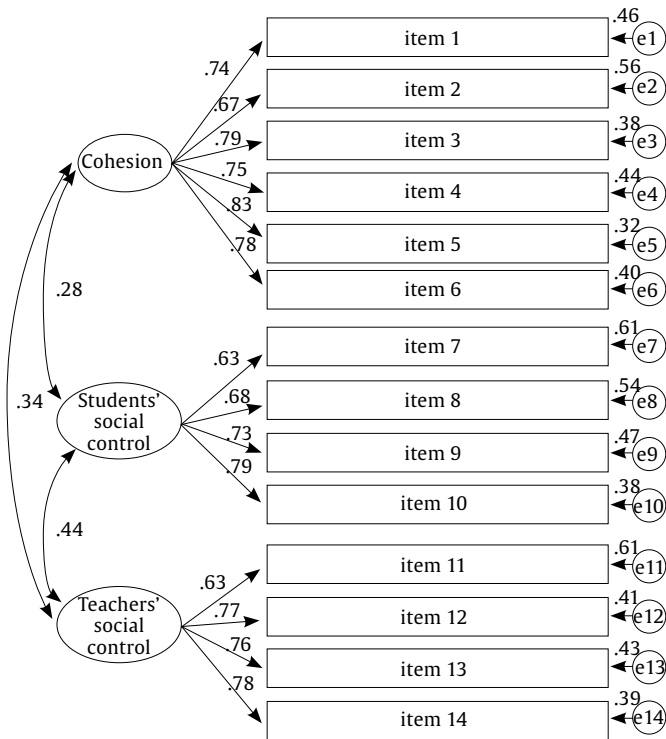


Figure 2. Final Factor Model of School Collective Efficacy to Bullying Prevention Depicting a Three-Dimensional First-Order Factor.

All correlations and standardized factor loadings are statistically significant ($p < .001$). See complete item questions in Appendix.

Assessing Measurement Invariance by Gender

As shown in Table 3, the goodness of fit statistic suggests a measurement invariance for this model by gender. The configural

model, which imposed no equality constraints between groups, served as a baseline model. Results showed the goodness of fit statistic had an acceptable fit to this model ($SB\chi^2 = 85.98$, $df = 60$, $p = .016$; $SRMR = .05$; $CFI = .98$; $TLI = .97$; $RMSEA = .03$, CI [.01, .04]). When all factor loadings were constrained to be equal across gender (metric invariance), the difference between this model and the configural model was not statistically significant, $\Delta SB\chi^2 = 5.67$, $df = 7$, $p = .578$, also the CFI and RMSEA values differed by no more than .01 and .015, respectively. In addition, intercepts of the observed variables were constrained to be equal by gender, with the differences of the metric model being not statistically significant, $\Delta SB\chi^2 = 20.29$, $df = 6$, $p = .317$, and differences in CFI and RMSEA are small (CFI < .01, RMSEA < .015).

Assessing Measurement Invariance by Educational Level (Secondary vs. High School)

To examine whether the three-dimensional factor model is invariant between secondary and high school students, a nested model similar to those explored for gender were tested (see Table 3). The baseline model (configural invariance) fit well to the data, which confirmed the validity of the model's internal structure ($SB\chi^2 = 103.11$, $df = 60$, $p < .001$; $SRMR = .05$; $TLI = .96$; $CFI = .97$; $RMSEA = .03$ CI 90 [.02, .04]). The model with factor loadings constrained (measure invariance) across groups fit well, based on the criteria of the $SB\chi^2$ difference and the change of CFI and RMSEA ($\Delta SB\chi^2 = 16.95$, $\Delta df = 12$, $p = .151$). These results suggest that the factor loadings are consistent for secondary and high school students. Finally, we added the intercept constraints to the model comparison (scalar invariance). In these models, there was no significant statistical differences in $SB\chi^2$ ($\Delta SB\chi^2 = 31.55$, $\Delta df = 6$, $p = .025$) and changes in CFI and RMSEA were smaller than the standard ($\Delta CFI = .004$, $\Delta RMSEA = .001$) suggested, so there are no substantial differences in the intercept of observed variables between early and middle adolescents.

Table 3. Summary of Fit Statistics for Testing Measurement Invariance of Three-Dimensional Model of School Collective Efficacy to Prevent Bullying ($N = 1,168$)

Invariance	$SB\chi^2$	df	$\Delta SB\chi^2$	Δdf	p	ΔCFI	$\Delta RMSEA$
Gender							
Configural	85.98	60			.016		
Metric	91.65	67	5.67	7	.578	.001	.002
Scalar	115.23	85	23.58	18	.169	.004	.001
Educational level							
Configural	103.11	60			< .001		
Metric	110.76	67	7.64	7	.365	.001	.002
Scalar	142.31	73	31.55	6	.025	.004	.011

Assessing Group Differences in Three-Dimensional Model Factor Means

To estimate the difference between the three-dimensional model means by gender, the boys were chosen as a reference group and their

Table 4. Means Differences by Gender and Educational level on SCEE

Variable	Factor	M_{Diff}	z-statistics	p	Cohen's d
Gender	Cohesion	.21	3.03	.021	0.18
	Students' social control	.05	0.32	.642	0.02
	Teachers' social control	.10	1.33	.184	0.08
Educational level	Cohesion	-.04	-0.44	.655	0.03
	Students' social control	.30	4.85	<.001	0.28
	Teachers' social control	.09	1.09	.275	0.06

$p < .001$.

latent means were set to zero. Then, the latent means of the girls informed the difference in factor means between the two groups (see Table 4). The test revealed statistically significant mean differences by gender on one of the factors. Particularly, girls had higher scores on cohesion. The gender differences in students' informal social control, and teachers' informal social control were not statistically significant.

In terms of latent means differences by educational level, we chose the secondary students' group as the reference group and estimated the latent mean of the middle adolescent group. Results showed one statistically significant means difference between secondary and high school students (see Table 4). High school students had a higher score on students' informal social control than secondary students. The differences in cohesion and teachers' informal social control were not statistically significant.

Discriminant Validity

Table 5 shows the correlations between the scales, the squares of those correlations (in parenthesis), and the average variance explained (AVE). The squared of the correlations (R^2) between the subscales of SCEE is lower than AVE. This result indicates evidence of discriminant validity, according to the standard criteria (see Fornell & Larcker, 1981; Hair et al, 2010).

Concurrent Validity

With respect to concurrent validity, the subscales of SCEE correlated significantly with construct measures in the study (see Table 5). As expected, the subscales of SCEE had a positive correlation to school fairness and sense of school community and a negative correlation to school collective moral disengagement. Moreover, the effect-size of these correlations has potential consequences. The correlation between teachers' social control and school collective moral disengagement had a small effect-size ($r > .10$), which indicates a potential consequence at the level of individual events. The effect-size of all other correlations ranged between medium ($r > .20$) and large ($r > .30$), which may indicate important short and long-term practical consequences (Funder & Ozer, 2019). In summary, these results provide evidence of concurrent validity of the SCE subscales.

Table 5. Correlations Between SCEE Subscales and External Constructs

	SCEE-C AVE = .58	SCEE-S AVE = .50	SCEE-T AVE = .54	SCMD	SF	SSC
SCE-C	-					
SCE-S	.28*** (.07)	-				
SCE-T	.34*** (.11)	.28*** (.07)	-			
SCMD	-.22*** (.05)	-.20*** (.04)	-.15*** (.02)	-		
SF	.38*** (.14)	.29*** (.08)	.25*** (.06)	-.36*** (.13)	-	
SSC	.37*** (.13)	.30*** (.09)	.39*** (.15)	-.20*** (.04)	.33*** (.11)	-

Note. SCEE-C = cohesion; SCEE-S = students' social control; SCEE-T = teachers' social control; SCMS = school collective moral disengagement; SF = school fairness; SSC = school sense of community. Square correlations are reported in parentheses (R^2).

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

Discussion

In the present study, uni-dimensional and multidimensional conceptualizations of school collective efficacy were tested. To do so, we examined psychometric properties of these models, as well as the invariance and external validity (discriminant and concurrent) of a three-dimensional factors model. Based on Williams and Guerra (2011), we hypothesized that the three-factor structure is better fit to measure school collective efficacy in the prevention of bullying. This study contributes to the literature regarding school collective efficacy in several ways. First, findings confirmed a better adjustment to the data of the three-factors model and found empirical evidence to support a multidimensional structure of the school collective efficacy construct. Second, empirical evidence supports the three-dimensional factor model for boys and girls, and secondary and high school students. This psychometric feature is critical for studies about antecedents and consequences of school collective efficacy. Finally, in terms of correlations between scales, we found expected correlations that provide evidence of discriminant and concurrent validity.

School Collective Efficacy as a Multidimensional Construct (Hypothesis 1 internal structure)

Findings support the theory that school collective efficacy is a multidimensional construct (Sampson et al., 1997; Sampson & Raudenbush, 1999). The hypothesized three-dimensional model showed that responses to the measurement of school collective efficacy can be grouped in three first-order factors (cohesion, students' social control, and teachers' social control). Compared to one-dimensional and two-dimensional factor models, the three-dimensional model had better adjustment to the data. Future research should examine whether the different dimensions of school collective efficacy have different antecedents and consequences.

Measurement Invariance across Gender and Educational Level (Hypothesis 2 measurement invariance)

The results provide empirical evidence supporting measurement invariance of school collective efficacy to prevent bullying by

gender and educational level (secondary vs. high school). For group comparisons by gender and education level, school collective efficacy can be measured with the same metric. Thus, score differences in the factors reflect group differences in the perception of school collective efficacy rather than item bias. Furthermore, we are confident that comparing variations of the construct in these groups are meaningful. Establishing measurement invariance allows research to make appropriate conclusions about antecedent and consequence variables associated with school collective efficacy to bullying prevention in these groups.

Means Comparison (Hypothesis 3 means comparison)

Considering that both factor loading and intercepts were invariant, we examined latent means differences on the three-dimensional model. Regarding gender differences, results suggest that mean scores for girls were higher than for boys in perception of cohesion in school, but not in students' social control or teachers' social control. These results are consistent with other recent research (Barchia & Bussey, 2011; Cho, 2017; Lee & Jo, 2017), and thus provide validity to these scales. With regard to educational level differences, we found the mean scores for high school students to be higher than secondary students on perceptions of students' social control but not on cohesion or teachers' social control. We suspect that these differences may reflect psychosocial developmental differences. Future research should examine causes and consequences of differences by gender and educational level in students' perception of school collective efficacy to prevent bullying.

Discriminant Validity (Hypothesis 4 discriminant validity)

In this study, based on Williams and Guerra (2011), we hypothesized that school collective efficacy to prevent bullying has a three-dimensional internal structure. Relationships between these factors should be further established to guide studies about the construct. Results about discriminant validity of the factors have shown that each subscale measures a subconstruct of a different nature (Barchia & Bussey, 2011; Capone et al., 2018; Hymel et al., 2015). These results suggest that research in bullying should analyze the causes and consequences of these dimensions in bullying prevention.

Concurrent Validity (Hypothesis 5 discriminant validity)

Moreover, the SCE showed concurrent validity. We found positive, statistically-significant correlations between school collective efficacy factors, school fairness, and sense of school community. Also, a negative correlation to school collective moral disengagement was found. The values of these correlations provide evidence of effect sizes with practical consequences in short- and long-term. In summary, findings suggest that these whole-school variables are associated and should be possible to consider them as indicators of a healthy school.

Practical Implications

The findings suggest that unidimensional interpretations for school collective efficacy to prevent bullying may not provide sufficient evidence of associated factors. In fact, this study provides empirical support to measure school collective efficacy as a three-dimensional construct. Our study provides a scale to practitioners for assessing students' perception of school collective efficacy to prevent bullying as an element to design school interventions able to reduce

bullying rates. Practitioners have a useful instrument to measure a set of factors of school collective efficacy that can be targeted in antibullying intervention. The analysis of means differences allows for identifying groups of students (boys and secondary students) who are perceived to have less healthy school functioning, and suggests that intervention should focus on them in particular.

Our analysis suggests school collective efficacy is associated to other indicators of whole-school health. Specifically, it shows that dimensions of school collective efficacy have positive associations to school fairness and sense of community and has a negative influence on school collective moral disengagement. As research indicates, school collective efficacy is associated with a healthy whole-school climate and bullying prevention (e.g., Goddard et al., 2015; Hymel et al., 2015; Olsson et al., 2017; Thornberg et al., 2019). Practitioners may consider aiming to intervene in the different dimensions of school collective efficacy studied here in order to enhance students' perceptions of school safety and bullying prevention. Students with lower views of school collective efficacy could be identified for special interventions. These prevention efforts are more effective in bullying prevention when they involve strategies that improve school members' cohesion, students' informal social control, and teachers' informal social control.

Limitations

Several limitations of this study should be taken into account. First, results rely on self-reported data aimed to assess school collective efficacy to bullying prevention. Therefore, we acknowledge further studies should include multiple sources of data (e.g., teachers and principals) and measurement methods (e.g., interview or observation) as well, as a means to offer a more robust instrument to measure school collective efficacy. Second, our findings are from a sample from a specific urban region of Mexico. It is essential to consider that adolescents from rural and indigenous communities might have different experiences regarding school efficacy to prevent bullying. It is desirable to use cross-national samples from different age groups and across diverse environments to provide empirical support for the school collective efficacy scale in future studies. Third, although school collective efficacy is a relevant variable to prevent bullying, more research is needed to establish an association between a new measure and bullying. Fourth, the same sample was used to establish the dimensionality of the scale and to study its measurement invariance, which in turn may be overestimating the fit of invariance models. Finally, the cross-sectional nature of the design prevents from assessing how the construct changes or remains constant across time. Future studies should be conducted to examine the changes in school collective efficacy across time.

Implications for Future Research

These findings have methodological implications for researchers studying issues relative to school collective efficacy to prevent bullying, as well as practitioners seeking to hinder bullying in schools that rely, in part, on the extent to which they can promote cohesion, students' social control, and teachers' social control. Given the values of the construct of school collective efficacy in bullying research, it is critical that studies of school collective efficacy use the best measurement instruments possible to ensure usable inferences. Future studies on the conceptualization and assessment of students' perception of school collective efficacy to bullying prevention should be enhanced by the theoretical argument and empirical results of this study. Additionally, the measurement invariance of factor structure is relevant for practitioners who implement whole-school bullying prevention programs who might want to assess whether

interventions have different effects on different groups within their school. Finally, the correlation of collective efficacy with positive whole-school variables implies that further studies should deepen the relationships between these variables.

Conflict of Interest

The authors of this article declare no conflict of interest.

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*(3), 488-499. <https://doi.org/10.1037/0022-006X.62.3.488>
- Albanesi, C., Cicognani, E., & Zani, B. (2007). Sense of community, civic engagement and social well-being in Italian adolescents. *Journal of Community & Applied Psychology, 17*(5), 387-406. <https://doi.org/10.1002/casp.903>
- Allison, K. P., & Bussey, K. (2017). Individual and collective moral influences on intervention in cyberbullying. *Computer in Human Behavior, 74*, 7-15. <https://doi.org/10.1016/j.chb.2017.04.019>
- Almanasreh, E., Moles, R., & Chen, T. F. (2019). Evaluation of methods used for estimating content validity. *Research in Social and Administrative Pharmacy, 15*(2), 214-221. <https://doi.org/10.1016/j.sapharm.2018.03.066>
- Bacon, D. R., Sauer, P. L., & Young, M. (1995). Composite reliability in structural equations modeling. *Educational and Psychological Measurement, 55*(3), 394-406. <https://doi.org/10.1177/0013164495055003003>
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Prentice-Hall.
- Bandura, A. (1989). Social cognitive theory. In R. Vasta (Ed.), *Annals of child development. Vol 6: Six theories of child development* (pp. 1-60). JAI Press.
- Bandura, A. (2000). Exercise of human through collective efficacy. *Current Directions in Psychological Science, 9*(3), 75-78. <https://doi.org/10.1111/1467-8721.00064>
- Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology, 52*(1), 1-26. <https://doi.org/10.1146/annurev.psych.52.1.1>
- Bandura, A. (2002). Selective moral disengagement in the exercise of moral agency. *Journal of Moral Education, 31*(2), 101-119. <https://doi.org/10.1080/0305724022014322>
- Barchia, K., & Bussey, K. (2011). Individual and collective social cognitive influences on peer aggression: Exploring the contribution of aggression efficacy, moral disengagement, and collective efficacy. *Aggressive Behavior, 37*(2), 107-120. <https://doi.org/10.1002/ab.20375>
- Black, D. (1984). Social control as a dependent variable. In D. Black (Ed.), *Studies on low and social control* (pp. 1-36). Academic Press. <https://doi.org/10.1016/B978-0-12-102801-5.50007-8>
- Bourdieu, P. (1984). *Distinction: A social critique of the judgment of taste*. Harvard University Press.
- Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). The Guilford Press.
- Byrne, B. M. (2016). *Structural equation modeling with Mplus. Basic concepts, applications, and programming* (4th ed.). Routledge.
- Capone, V., Donizzeti, A. R., & Petrillo, G. (2018). Classroom relationships, sense of community, perceptions of justice, and collective efficacy for students' social well-being. *Journal of Community Psychology, 46*(3), 374-382. <https://doi.org/10.1002/jcop.21943>
- Cheung, G. W., & Rensvold, R. B. (2002) Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9*(2), 233-255. https://doi.org/10.1207/S15328007SEM0902_5
- Cho, S. (2017). Self-control and risky lifestyles in context: Cross-level integration between opportunity and collective efficacy in the study of peer victimization among South Korean youth. *Journal of Child and Family Studies, 26*(1), 67-79. <https://doi.org/10.1007/s10826-016-0554-y>
- Coleman, J. S. (1990). *Foundations of social theory*. Harvard University Press.
- de Oliveira W. A., Silva M. A., da Silva J. L., de Mello F. C., do Prado R. R., & Malta D. C. (2015). Associations between the practice of bullying and individual and contextual variables from the aggressors' perspective. *Journal de Pediatria, 92*(1), 32-39. <https://doi.org/10.1016/j.jped.2015.04.003>
- du Plessis, G. A., & de Bruin, G. P. (2015). Using Rasch modelling to examine the international personality item pool (IPIP) values in action (VIA) measure of character strengths. *Journal of Psychology in Africa, 25*(6), 512-521. <https://doi.org/10.1080/14330237.2015.1124603>
- du Plessis, M. R., Smeekens, S., Cillessen, A. H. N., Whittle, S., & Guroglu, B. (2019). Bullying the brain? Longitudinal links between childhood peer victimization cortisol, and adolescent brain structure. *Frontiers in Psychology, 9*, e2706. <https://doi.org/10.3389/fpsyg.2018.02706>
- Espegale, D. L., Rao, M. A., & De La Rue, L. (2013). Current research on school-based bullying: A social-ecological perspective. *Journal of Social Distress and Homelessness, 22*(1), 21-27. <https://doi.org/10.1179/1053078913Z.0000000002>
- Farrell, A. M. (2010). Insufficient discriminant validity: A comment on Bove, Pervan, Beatty, and Shiu (2009). *Journal of Business Research, 63*(3), 324-327. <https://doi.org/10.1016/j.jbusres.2009.05.003>
- Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*(1), 39-50. <https://doi.org/10.1177/002224378101800104>
- Funder, D. C., & Ozer, D. J. (2019). Evaluating effect size in psychological research: Sense and nonsense. *Advances in Methods and Practices in Psychological Science, 2*(2), 156-168. <https://doi.org/10.1177/2515245919847202>
- Furr, R. M., & Bacharach, V. R. (2014). *Psychometrics. An introduction* (2nd ed.). Sage.
- Garaigordobil, M., & Machimbarrena, J. M. (2019). Victimization and perpetration of bullying/cyberbullying: Connections with emotional and behavioral problems and childhood stress. *Psychosocial Intervention, 28*(2), 67-73. <https://doi.org/10.5093/pi2019a3>
- Gaskin, C. J., & Happell, B. (2014). On exploratory factor analysis: A review of recent evidence, an assessment of current practice, and recommendations for future use. *International Journal of Nursing Studies, 51*(3), 511-521. <https://doi.org/10.1016/j.ijnurstu.2013.10.005>
- Gini, G., Pozzoli, T., & Bussey, K. (2015). The role of individual and collective moral disengagement in peer aggression and bystanding: A multilevel analysis. *Journal of Abnormal Child Psychology, 43*, 441-452. <https://doi.org/10.1007/s10802-014-9920-7>
- Goddard, R., Goddard, Y., Kim, E. S., & Miller, R. (2015). A theoretical and empirical analysis of the roles of instructional leadership, teacher collaboration, and collective efficacy belief in support of student learning. *American Journal of Education, 121*(4), 501-530. <https://doi.org/10.1086/681925>
- Green, S. B. (2015). Evaluation of dimensionality in the assessment of internal consistency reliability: Coefficient alpha and omega coefficients. *Educational Measurement: Issues and Practice, 34*(4), 14-20. <https://doi.org/10.1111/emip.12100>
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis* (7th ed.). Pearson.
- Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2017). *A primer on partial least squares structural equation modeling* (2nd ed.). Sage.
- Häuberer, J. (2011). The founding concepts of social capital – Bourdieu's theory of capital and Coleman's rational-choice approach to social capital. In J. Häuberer (Ed.), *Social capital theory. Towards a methodological foundation* (pp. 35-51). Springer.
- Hirschi, T. (2001). On the compatibility of rational choice, and social control theories of crime. In D. B. Cornish, & R. V. Clarke (Eds.), *The reasoning criminal. Rational choice perspectives on offending* (pp. 105-118). Transaction Publishers.
- Hu, L.-T., Bentler, B. M., & Kano, Y. (1992). Can test statistic in covariance structure analysis be trusted? *Psychological Bulletin, 112*(2), 351-362. <https://doi.org/10.1037/0033-2909.112.2.351>
- Hymel, S., McClure, R., Miller, M., Shumka, E., & Trach, J. (2015). Addressing school bullying: Insights from theories of group processes. *Journal of Applied Developmental Psychology, 37*, 16-24. <https://doi.org/10.1016/j.appdev.2014.11.008>
- Instituto Nacional para la Evaluación de la Educación (INEE, 2018). <https://historico.mejoredu.gob.mx/00-inicio/instituto-nacional-para-la-evaluacion-de-la-educacion/>
- Janowitz, M. (1975). Sociological theory and social control. *American Journal of Sociology, 81*(1), 82-108. <https://doi.org/10.1086/226035>
- Kline, R. (2013). Exploratory and confirmatory factor analysis. In Y. Petscher, C. Schatschneider, & D. L. Compton (Eds.), *Applied quantitative analysis in education and the social sciences* (pp. 171-207). Routledge.
- Kollerová, L., Soukup, P., & Gini, G. (2018). Classroom Collective Moral Disengagement Scale: Validation in Czech adolescents. *European Journal of Developmental Psychology, 15*(2), 184-191. <https://doi.org/10.1080/17405629.2017.1292907>
- Konishi, C., Miyazaki, Y., Hymel, S., & Waterhouse, T. (2017). Investigating associations between school climate and bullying in secondary schools: Multilevel contextual effects modeling. *School Psychology International, 38*(3), 240-263. <https://doi.org/10.1177/0143034316688730>
- Kurnianingshi, S., Yunarti, K. W., Kim, U. (2012). Factors influencing trust of teachers among students. *International Journal of Research Studies in Education, 1*(2), 85-94. <https://doi.org/10.5861/ijrse.2012.v1i2.77>
- Lee, B., & Jo, Y. (2017). Correlates of repeat victimization and sex differences in South Korean youth. *Journal of Interpersonal Violence, 35*(17-18), 3188-3216. <https://doi.org/10.1177/08862605170708760>
- Lenzi, M., Vieno, A., Gini, G., Pozzoli, T., Pastore, M., Santinello, M., & Elgar, F. J. (2014). Perceived teacher unfairness, instrumental goals, and bullying behavior in early adolescence. *Journal of Interpersonal Violence, 29*(10), 1834-1849. <https://doi.org/10.1177/0886260513511694>
- Lynn, M. R. (1986). Determination and quantification of content validity. *Nursing Research, 35*(6), 382-386. <https://doi.org/10.1097/00006199-198611000-00017>

- Marées, N. V., & Petermann, F. (2010). Bullying in German primary schools: Gender differences, age trends and influence of parents' migration and educational backgrounds. *School Psychology International*, 31(2), 178-198. <https://doi.org/10.1177/0143034309352416>
- Martínez-Ferrer, B., Moreno-Ruiz, D., Amador, L. V., & Orford, J. (2011). School victimization among adolescents. An analysis from an ecological perspective. *Psychosocial Intervention*, 20(2), 149-160. <https://doi.org/10.5093/in2011v20n2a3>
- Millsap, R. E., & Olivera-Aguilar, M. (2012). Investigating measurement invariance using confirmatory factor analysis. In R. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 380-416). The Guilford Press.
- Modecki, K. L., Minchin, J., Harbaugh, A. G., Guerra, N. G., & Runions, K. C. (2014). Bullying prevalence across contexts: A meta-analysis measuring cyber and traditional bullying. *Journal of Adolescent Health*, 55(5), 602-611. <https://doi.org/10.1016/j.jadohealth.2014.06.007>
- Muthén, L. K., & Muthén, B. O. (2017). *Mplus user's guide* (8th ed.). Muthén & Muthén.
- Nielsen, M. B., Tangen, T., Idsoe, T., Matthiesen, S. B., & Magerøy, N. (2015). Post-traumatic stress disorder as a consequence of bullying at work and at school. A literature review and meta-analysis. *Aggression and Violent Behavior*, 21, 17-24. <https://doi.org/10.1016/j.avb.2015.01.001>
- Olsson, G., Laftman, S. B., & Modin, B. (2017). School collective efficacy and bullying behaviour: A multilevel study. *International Journal of Environmental Research and Public Health*, 14(12), e1607. <https://doi.org/10.3390/ijerph14121607>
- Olweus, D. (1993). *Bullying at school*. Blackwell Publishing.
- Park, H. S., Dailey, R., & Lemus, D. (2002). The use of exploratory factor analysis and principal components analysis in communication research. *Human Communication Research*, 28(4), 562-577. <https://doi.org/10.1111/j.1468-2958.2002.tb00824.x>
- Payne, A. A., & Gottfredson, D. C. (2004). School and bullying: School factors related to bullying and school-based bullying interventions. In C. E. Sanders, & G. D. Phye (Eds.), *Bullying. Implications for the classroom* (pp. 159-176). Academic Press.
- Peguro, A. A., Bondy, J. M., & Shekarkhar, Z. (2017). Punishing Latino/a youth: School justice, fairness, order, dropping out, and gender disparities. *Hispanic Journal of Behavioral Sciences*, 39(1), 98-125. <https://doi.org/10.1177/0739986316679633>
- Peterson, C., & Seligman, M. (2004). *Character strengths and virtues: A handbook and classification*. Oxford University Press.
- Pozzoli, T., Gini, G., & Vieno, A. (2012). The role of individual correlates and class norms in defending and passive bystander behavior in bullying: A multilevel analysis. *Child Development*, 83(6), 1917-1931. <https://doi.org/10.1111/j.1467-8624.2012.01831.x>
- Prati, G., Cicognani, E., & Albanesi, C. (2017). Psychometric properties of multidimensional scale of sense of community in the school. *Frontiers in Psychology*, 8, e1488. <https://doi.org/10.3389/fpsyg.2017.01466>
- Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review*, 41, 71-90. <https://doi.org/10.1016/j.dr.2016.06.004>
- Raftery, A. E. (1995). Bayesian model selection in school research. *Sociological Methodology*, 25, 111-163. <https://doi.org/10.2307/271063>
- Roland, E., & Idsoe, T. (2001). Aggression and bullying. *Aggressive Behavior*, 27(6), 446-462. <https://doi.org/10.1002/ab.1029>
- Sampson, R. J. (2006). Collective efficacy theory: Lesson learned and directions for future inquiry. In F. T. Cullen, J. P. Wright, & R. Blevins (Eds.), *Advances in criminological theory*. Vol. 15: *Taking stock. The status of criminology theory* (pp. 149-167). Transaction Publishers.
- Sampson, R. J., Morenoff, J. D., & Earls, F. (1999). Beyond social capital: Spatial dynamics of collective efficacy for children. *American Sociological Review*, 64(5), 633-660. <https://doi.org/10.2307/2657367>
- Sampson, R. J., & Raudenbush, S. W. (1999). Systematic social observation of public spaces: A new look at disorder in urban neighborhoods. *American Journal of Sociology*, 105(3), 603-651. <https://doi.org/10.1086/210356>
- Sampson, R. J., Raudenbush, S. W., & Earls, F. (1997). Neighborhoods and violent crime: A multilevel study of collective efficacy. *Science*, 277(5328), 918-924. <https://doi.org/10.1126/science.277.5328.918>
- Sapouna, M. (2010). Collective efficacy in the school context: Does it help explain victimization and bullying among Greek primary and secondary school students? *Journal of Interpersonal Violence*, 25(10), 1912-1927. <https://doi.org/10.1177/0886260509354509>
- Sass, D. A., & Schmitt, T. A. (2013). Testing measurement and structural invariance. In T. Teo (Eds.), *Handbook of quantitative methods for educational research* (pp. 315-345). Sense Publishers.
- Satorra, A., & Bentler, P. M. (2010). Ensuring positiveness of the scaled difference chi-square test statistic. *Psychometrika*, 75(2), 243-248. <https://doi.org/10.1007/s11336-009-9135-y>
- Sharma, S., Mukherjee, S., Kumar, A., & Dillon, W. R. (2005). A simulation study to investigate the use of cutoff values for assessing model fit in covariance structure models. *Journal of Business Research*, 58(7), 935-943. <https://doi.org/10.1016/j.jbusres.2003.01.007>
- Sigurdson, J. F., Undheim, A. M., Wallander, J. L., Lydersen, S., & Sund, A. M. (2015). The long-term effects of being bullied or a bully in adolescence on externalizing and internalizing mental health problems in adulthood. *Child & Adolescent Psychiatry & Mental Health*, 9, e42. <https://doi.org/10.1186/s13034-015-0075-2>
- Sourander, A., Lempinen, L., & Brunstein, A. (2016). Changes in mental health, bullying behavior, and service use among eight-year-old children over 24 years. *Journal of the American Academy of Child & Adolescent Psychiatry*, 55(8), 717-725. <https://doi.org/10.1016/j.jaac.2016.05.018>
- Swearer, S. M., & Hymel, S. (2015). Understanding the psychology of bullying. Moving towards a social-ecological diathesis-stress model. *American Psychologist*, 70(4), 344-353. <https://doi.org/10.1037/a0038929>
- Taylor, C. (2001). The relationship between social and self-control: Tracing Hirschi's criminological career. *Theoretical Criminology*, 5(3), 369-388. <https://doi.org/10.1177/1362480601005003004>
- Thornberg, R., & Elvstrand, H. (2012). Children's experiences of democracy, participation, and trust in school. *International Journal of Educational Research*, 53, 44-54. <https://doi.org/10.1016/j.ijer.2011.12.010>
- Thornberg, R., Wänström, L., Elmeli, R., Johansson, A., & Mellander, E. (2020). Standing up for the victim or supporting the bully? Bystander responses and their associations with moral disengagement, defender self-efficacy, and collective efficacy. *Social Psychology of Education*, 23, 563-581. <https://doi.org/10.1007/s11218-020-09549-z>
- Thornberg, R., Wänström, L., & Hymel, S. (2019). Individual and classroom social-cognitive processes in bullying: A short-term longitudinal multilevel study. *Frontiers in Psychology*, 10, e1752. <https://doi.org/10.3389/fpsyg.2019.01752>
- Tomarken, A. J., & Waller, N. G. (2005). Structural equation modeling: Strengths, limitations, and misconceptions. *Annual Review of Clinical Psychology*, 1, 31-65. <https://doi.org/10.1146/annurev.clinpsy.1.102803.144239>
- Trach, J., Hymel, S., & Shumka, E. (2013). *Examining the influence of moral disengagement and collective efficacy on bullying: A multilevel analysis*. Paper presented at the biennial meeting of the Society for Research in Child Development, Seattle WA.
- United Nations Educational, Scientific, and Cultural Organization (UNESCO, 2017). *School violence and bullying. Global status report*. UNESCO.
- Valdés-Cuervo, A. A., Martínez-Ferrer, B., & Carlos-Martínez, E. A. (2018). The role of teaching practices in the prevention of school violence among peers. *Revista de Psicodidáctica*, 23(1), 33-38. <https://doi.org/10.1016/j.psicod.2017.05.006>
- Van de Schoot, R., Schmidt, P., De Beuckelaer, A., Lek, K., & Zondervan-Zwijnenberg, M. (2015). Editorial: Measurement invariance. *Frontiers in Psychology*, 6, e1064. <https://doi.org/10.3389/fpsyg.2015.01064>
- Vieno, A., Lenzi, M., Santinello, M., & Scacchi, L. (2013). Sense of community, unfairness, and psychosomatic symptoms: A multilevel analysis of Italian schools. *Journal of Adolescent Health*, 53(1), 142-145. <https://doi.org/10.1016/j.jadohealth.2013.02.019>
- Volk, A. A., Dane, A. V., & Marini, Z. A. (2014). What is bullying? A theoretical re-definition. *Development Review*, 34(4), 327-343. <https://doi.org/10.1016/j.dr.2014.09.001>
- Wach, S., Bilz, L., Fischer, S. M., Schubarth, W., & Wright, M. F. (2018). Students' willingness to intervene in bullying: Direct and indirect associations with classroom cohesion and self-efficacy. *International Journal of Environmental and Public Health*, 15, e2577. <https://doi.org/10.3390/ijerph15112577>
- Wang, J., & Wang, X. (2012). *Structural equation modeling applications using Mplus*. John Wiley & Sons.
- Wänström, L., Pozzoli, T., Gini, G., Thornberg, R., & Alsaadi, S. (2017). Perceived collective efficacy to stop aggression at school: A validation of an Italian and Swedish version of a scale for adolescents. *European Journal of Developmental Psychology*, 16(3), 349-361. <https://doi.org/10.1080/17405629.2017.1414695>
- Wiatrowski, M. D., Griswold, D. B., & Roberts, M. K. (1981). Social control theory and delinquency. *American Psychological Review*, 46(5), 525-541. <https://doi.org/10.2307/2094936>
- Widaman, K. F. (1993). Common factor analysis versus principal component analysis: Differential bias in representing model parameters? *Multivariate Behavioral Research*, 28(3), 263-311. https://doi.org/10.1207/s15327906mbr2803_1
- Widaman, K. F., & Resie, S. P. (1997). Exploring the measurement invariance of psychological instruments: Applications in the substance use domain. In K. J. Bryant, M. Windle, & S. G. West (Eds.), *The science of prevention: Methodological advances from alcohol and substance abuse research* (pp. 281-324). American Psychological Association.
- Wilcox, P., Cullen, F. T., & Feldmeyer, B. (2018). *Communities and crime. An enduring American challenge*. Temple University Press.
- Williams, K., & Guerra, N. (2011). Perceptions of collective efficacy and bullying perpetration in schools. *Social Problems*, 58(1), 126-143. <https://doi.org/10.1525/sp.2011.58.1.126>
- Zych, I., Farrington, D. P., & Tfofi, M. M. (2017). School bullying in different countries: Prevalence, risk factors, and short-term outcomes. In I. Zych, D. P. Farrington, V. J. Llorent, & M. M. Tfofi (Eds.), *Protecting children against bullying and its consequences* (pp. 5-22). Springer.

Appendix

Measurement Models of School Collective Efficacy

Items	One-factor	Two-factor		Three-factor		
		C	SC	C	SSC	TSC
<i>I believe my school has achieved...</i>	✓	✓		✓		
1. To build trust between students	✓	✓		✓		
2. Respectful relationships between students	✓	✓		✓		
3. Agreement from students to end bullying	✓	✓		✓		
4. Agreement from teachers to end bullying	✓	✓		✓		
5. That everyone takes care of each other	✓	✓		✓		
6. That school members share antibullying beliefs	✓	✓		✓		
<i>Students in my school stop bullying when...</i>						
7. A student is victim of teasing from a stronger student	✓		✓		✓	
8. A student is victim of rumors and lies	✓		✓		✓	
9. A student is being teased on the internet	✓		✓		✓	
10. A student is excluded by others	✓		✓		✓	
<i>Teachers in my school stop bullying when...</i>						
11. A student is victim of teasing from a stronger student	✓		✓			✓
12. A student is victim of rumors and lies	✓		✓			✓
13. A student is being teased on the internet	✓		✓			✓
14. A student is excluded by others	✓		✓			✓

Note. Checkmark indicates that item is included in the factor; C = cohesion; SC = informal social control; SSC = students' informal social control; TSC = teachers' informal social control.