

Brief Report

The Personal Experiences Checklist Short Form (PECK-SF):

Tests of invariance across gender and over time.

Dr Nathalie Noret^{1*} Professor Simon C. Hunter^{2,3} & Dr Susan Rasmussen⁴

Authors note

Supplementary materials, including a copy of analysis scripts and data files can be found at: <https://doi.org/10.17605/OSF.IO/Q7C98> . Data used in this research study were not collected as part of any funded project. This study was not pre-registered. However, all data have been made openly available at the Open Science Framework and can be accessed at <https://osf.io/q7c98/>. NN, SH, SR contributed to the conceptualisation of the study. NN gathered, curated and analysed the data. NN wrote the original draft. SH and SR provided supervision and reviewed and edited the manuscript.

*Corresponding author

¹Department of Education, University of York, York YO10 5DD UK

²Department of Psychology, Glasgow Caledonian University, Cowcaddens Road, Glasgow, G4 0BA UK

³Graduate School of Education, The University of Western Australia, WA, Australia

⁴School of Psychological Sciences and Health, University of Strathclyde, Graham Hills Building, 40 George Street, Glasgow, G1 1QE

PECK-SF: Tests of invariance across gender and over time.

Abstract

Peer-victimization is a frequent experience for many children and adolescents. Accurate measurement of peer-victimization is essential for better understanding such experiences and informing intervention work. While many peer-victimization and bully scales exist, they often lack important psychometric information. The short form of the Personal Experiences Checklist (PECK-SF) provides a brief measure of peer-victimization. Existing psychometric data supports the use of the scale as a measure of peer-victimization. However, there are no reports of invariance testing, limiting our ability to confidently use the scale to compare scores across boys and girls and for use in longitudinal research studies. Therefore, the aim of this study is to test the PECK-SF for invariance across gender and over time. Data were collected as part of a longitudinal study of 11- to 14-year-olds who provided data at two-time points two months apart ($N_{\text{Time 1}}=744$ $N_{\text{Time 2}}=333$). Confirmatory factor analysis supported the one-factor structure of the PECK-SF. Invariance testing demonstrated strict invariance over time and across gender. The use of the PECK-SF is recommended when examining gender differences, or conducting longitudinal studies, in peer-victimization.

Public significance statement

The results of this study highlight that the PECK-SF provides a short measure of experiences of peer-victimization appropriate for testing gender differences and changes over time.

Keywords

Bullying, Peer-Victimisation, Confirmatory Factor Analysis (CFA), Measurement Invariance

PECK-SF: Tests of invariance across gender and over time.

Despite the proliferation of anti-bullying programmes (Gaffney et al., 2021), peer-victimization continues to be a frequent experience for many children and adolescents. Experiencing peer-victimization in childhood and adolescence is related to a range of negative outcomes, including poor mental health, poorer academic achievement, and anti-social behavior (Gardella et al., 2017; Reijntjes et al., 2010; Valdebenito et al., 2017). Given the impact and prevalence of peer-victimization, it is critical that well-validated and reliable measurement scales are available to investigate experiences of peer-victimization and evaluate anti-bullying interventions. While many bullying and peer-victimization scales have been developed, they often lack clear psychometric data and, therefore, can be described as being in the early stages of development (Volk et al., 2017). The short form of the Personal Experiences Checklist (PECK-SF) (Prinz et al., 2019) provides a general, brief 14-item measure of peer-victimization experiences. While early evidence suggests the PECK-SF is a valid and reliable measure, further evidence as to the psychometric properties and, particularly, invariance of the scale is needed. Therefore, the aim of this study is to estimate the measurement invariance of the PECK-SF across gender and over time.

Peer-Victimisation

Peer-victimisation is defined as frequently experienced aggressive behavior within the peer group (Hunter et al., 2007). Peer-victimisation includes a range of aggressive behaviors, including direct verbal and physical acts such as name-calling and being hit or kicked, indirect acts such as being ignored or having rumours spread, and cyber-victimization such as being humiliated or embarrassed on social media and being sent nasty or threatening text messages. Approximately 25% of adolescents frequently experience such behaviors (Juvonen & Graham, 2014), and direct and indirect behaviors are more frequently experienced than cyber-aggression (Modecki et al., 2014). Bullying is a specific form of peer-victimization, which includes the additional characteristics of an imbalance of power between the

PECK-SF: Tests of invariance across gender and over time.

perpetrator and the target, and an intent to cause harm (Olweus, 1993). The definition of peer-victimization, however, omits any reference to power imbalance or intentionality and instead is defined as frequently experienced aggressive behaviour that occurs within the peer-group (Hunter, Boyle, & Warden, 2007; Söderberg & Björkqvist, 2020). The distinction between peer-victimization and bullying is often confused in the literature as the terms are frequently used interchangeably (Hunter et al., 2007). This definitional confusion has implications for the way peer-victimization and bullying are measured.

Assessment of peer-victimization and bullying using questionnaires raises two key methodological issues (Hunter et al., 2021). First, is the decision concerning whether a definitional or a behavioral approach should be taken (Vessey et al., 2014; Volk et al., 2017). The former presents participants with a definition of bullying explaining the critical characteristics of bullying before asking respondents how often they have experienced bullying behaviors within a given time frame (e.g., The Olweus Bully/Victim Questionnaire: Olweus, 1993). In contrast, and due to the potential for subjective interpretation when reading definitions of bullying, behavioral measures of bullying and peer-victimization typically omit any definition (e.g., the PECK-SF, Prinz et al., 2019). Instead, behavioral measures present participants with a list of aggressive behaviors and ask them to rate how often they have experienced each type of behavior. This conceptualizes experiences of bullying and peer-victimization as a continuum (Hunt et al., 2012).

The Personal Experiences Checklist

The Personal Experiences Checklist (PECK) (Hunt et al., 2012) is a well-validated 32-item measure of the frequency of peer-victimization experiences. Items were developed based on a literature search on bullying, discussion with experts in the field, and responses to open questions from previous survey work (Hunt et al., 2012). The scale captures four forms of peer-victimization: verbal-relational, physical, and cyberbullying behaviors, alongside

PECK-SF: Tests of invariance across gender and over time.

behaviors perpetrated because of culture. Participants are asked how often they have experienced each of the behaviors in the previous month, and responses are rated on a five-point Likert scale, ranging from “never” to “most days”. Hunt et al. (2012) report that the internal consistency of all four sub-scales was excellent ($\alpha=.91$ for verbal-relational bullying, $\alpha=.90$ for cyberbullying, $\alpha=.91$ for physical bullying, and $\alpha=.78$ for bullying based on culture). The four-factor structure was supported in both exploratory and confirmatory factor analyses, and test-retest reliability after two weeks was acceptable ($r = .75$ for relational-verbal bullying, $r = .86$ for cyberbullying, $r = .61$ for physical bullying, $r = .77$ for bullying based on culture, and $r = .79$ for the total PECK scale). Therefore, findings from the initial validation of the PECK suggest it is a psychometrically sound measure of peer-victimization (Hunt et al., 2012).

The short version of the Personal Experiences Checklist (PECK-SF) was developed in 2019 to provide a more rapid assessment tool with strong psychometric properties (Prinz et al., 2019). The PECK-SF was developed and validated with a pooled sample of over 1,000 children and adolescents drawn from two separate studies. The final 14 items were identified based on the content of the items and the size of the discrimination parameters calculated through the CFA. The initial analysis identified problematic item thresholds (distance between points on the ordinal response scale) between the ‘most days’ and ‘everyday’ categories. Therefore, these final points on the Likert scale were combined into a single category (most days / every day), changing the scale from a five-point to a four-point Likert scale. The subsequent analyses were calculated on the recoded data (Prinz et al., 2019). Regarding the structure of the scale, the results of the confirmatory factor analysis (CFA) supported the four-factor structure of the PECK-SF. However, the reliability of these subscales was poor. However, the one-factor structure of the scale was supported by CFA, and the reliability of this total PECK-SF scale was good ($\alpha=.84$). These findings support

PECK-SF: Tests of invariance across gender and over time.

using the PECK-SF as a unidimensional measure with a single, total score (Prinz et al., 2019).

The current study

Measures of peer-victimization are argued to be in the early stages of development due to the limited available evidence of the psychometric properties of such scales (Vessey et al., 2014; Volk et al., 2017). Without psychometrically strong measures, it is difficult to confidently identify and compare prevalence rates and effectively evaluate anti-bullying interventions (Volk et al., 2017). Invariance (equivalence) testing allows identification of whether the psychometric properties of a scale are similar across conditions (e.g., over time, across gender) (Putnick & Bornstein, 2016). This, in turn, implies that a measurement assesses the same thing, in the same way, across conditions or groups (Putnick & Bornstein, 2016) and that variation in scores reflects genuine differences across conditions and is not reflective of any psychometric artefacts of the scale (Cheung & Rensvold, 2002). However, the invariance of bullying or peer-victimization measures across conditions is not frequently reported (Vessey et al., 2014).

Much research on peer-victimization and bullying focuses on understanding experiences of these behaviors across different conditions or groups, for example, across gender and over time. Gender differences are often reported in experiences of bullying (e.g., Kennedy, 2019), reflecting possible differences in socialization experiences and different group norms regarding behavior (Smith et al., 2019). Furthermore, longitudinal investigations are an increasingly important focus of research on peer-victimization. Such studies typically examine the prevalence and impact of peer-victimization over time (e.g., Zych et al., 2020) or seek to evaluate anti-bullying programmes (e.g., Olweus et al., 2019). Ensuring peer-victimization scales are invariant across conditions, such as across gender and over time, ensures that any reported differences do not simply reflect artefacts of the scales.

PECK-SF: Tests of invariance across gender and over time.

When choosing a measure of peer-victimization or bullying, researchers and clinicians should consider the psychometric properties of the scale and ensure that the behavior being measured (bullying or peer-victimisation) is appropriate for its intended purpose. Despite the development of many measures of peer-victimization and bullying, many lack strong evidence on the psychometric properties of the scale, and many fail to make the relationship between their definition and their measure explicit (Vessey et al., 2014). Developed in 2019 (Prinz et al., 2019), the PECK-SF is a relatively new scale designed to provide a quick and short measure of the frequency of peer-victimization experiences. Initial psychometric data supports the use of the PECK-SF in providing a single score of peer-victimization experiences (Prinz et al., 2019). However, invariance testing of the measure across different conditions is currently lacking. Therefore, the aim of the current study is to assess the extent to which the PECK-SF is invariant across gender and time.

Method

Data for the study were collected as part of a two-month longitudinal investigation into the association between peer-victimization and depressive symptomology (reference withheld for review). This report includes data from participants who completed the survey at time 1 (N=744) and two months later at time 2 (N=333). Participants were between 11 and 14 years old (Mean age at time 1 = 12.72; SD = 0.84) and were recruited from four secondary schools (High Schools) in England and Scotland. At time 1, 342 (46%) were male, 366 (49.2%) were female, 28 (3.8%) reported that they preferred not to say, and 8 (1.1%) did not answer the question.

The PECK-SF was administered at the start of the project and again two months later to measure experiences of peer-victimization. The PECK-SF includes 14 items (e.g., “Other kids say mean things behind my back”), and participants were instructed: “Thinking about the last month or so at school, how often do the following things happen to you?” and were

PECK-SF: Tests of invariance across gender and over time.

then asked to rate each of the 14 items on a five-point frequency Likert scale ranging from 0 = “never” to 4 = “every day”.

Ethical approval for the study was granted by the university research ethics committee of the third author. Schools were contacted with an invitation to participate in the study. Once approval had been granted, parental consent was obtained. The survey was conducted during class time in the presence of a researcher. Participants were provided with the opportunity to consent to participate and reminded that they could withdraw at any time. Participants at two schools (N = 133) completed the survey online using the Qualtrics.com survey tool, participants at the other two schools (N = 928) completed a paper version of the survey. At the end of the survey, participants were provided with a debrief sheet with support and guidance information. This study was not pre-registered. However, all data have been made openly available at the Open Science Framework and can be accessed at (reference withheld for review).

Data analysis

Invariance testing was conducted using Mplus (Version 8.4) (Muthen & Muthen, 1998). At time 1, item level missing data varied between 1.7% (N=13) and 2.4% (N=17) of cases being missing. Therefore, missing data were treated using Full Information Likelihood Testing (FIML). The invariance testing was conducted across gender (for those identifying as male or female) and across time (time 1 and time 2). Items on the PECK-SF are responded to on a five-item ordered categorical scale from 0 to 4. Data are argued to be non-normal if skewness values exceed 2 and Kurtosis values exceed 7 (Curran et al, 1996). In the current study, at time 1, nine items had a skewness value that exceed 2 (ranging from 2.15 to 6.17) and six items had kurtosis values exceeding 7 (ranging from 7.62 to 41.62). Therefore, all paths were estimated using the weighted least square mean and variance

PECK-SF: Tests of invariance across gender and over time.

adjusted (WLSMV) estimator, which is the more appropriate estimator when analysing ordered categorical and non-normally distributed data (Sass et al., 2014).

Invariance testing was conducted in a CFA framework and followed a four-step process outlined by Bowen and Masa (2015). While this four-step process was initially developed for analysis using Maximum Likelihood (ML) estimation, evidence supports its use with the WLSMV estimator (Coertjens et al., 2012). Invariance was determined by analyzing changes in fit indices at each step. To support invariance, the difference in the comparative fit index (CFI) at each step should be less than 0.01, and the change in the root mean square error of approximation (RMSEA) should be below 0.015 (Cheung & Rensvold, 2002). The RMSEA point estimate of the preceding model should be included in the 90% confidence intervals of the new model. As chi-square tests are sensitive to sample size, they will not be reported (Cheung & Rensvold, 2002).

The reliability of the PECK-SF was also assessed across these conditions using McDonald's Omega. In comparison to Cronbach's alpha which can under-estimate reliability, McDonald's Omega does not assume tau-equivalence (equal factor loadings) or uncorrelated error variances (Hayes & Coutts, 2020; Trizano-Hermosilla & Alvarado, 2016). As such, McDonald's Omega is seen as a more robust indicator of reliability (Trizano-Hermosilla & Alvarado, 2016). McDonald's Omega is recommended as a measure of reliability for peer-victimisation data, due to the often skewed nature of data on the frequency of peer-victimisation experiences (Hunter et al., 2021). In the current study, McDonald's Omega was calculated to test the reliability of the PECK-SF for male and female participants, and across the two time points using Mplus (Version 8.4) (Muthen & Muthen, 1998), following the guidance provided by Hayes and Coutts (2020).

PECK-SF: Tests of invariance across gender and over time.

Results

Confirmatory factor analysis indicated that the model was an acceptable fit to the data χ^2 (df = 77) = 515.57, $p < .001$; CFI = .941; RMSEA = .088 (90% Confidence Interval = .081, .095). However, invariance testing for this model could not be calculated due to problems with the thresholds for items 6, 8, and 9, as there was a low frequency of responses for the final points of the Likert scale (“Most days” [4] and “Everyday” [5]). Therefore, and consistent with the approach taken by Prinz et al. (2019), the final points on the Likert scale were combined to create a four-point Likert scale ranging from never (0) to most days/ everyday (4). Fit assessed using the revised scale points was still an acceptable fit to the data: χ^2 (df = 77) = 487.17, $p < .001$; CFI = .934; RMSEA = .085 (90% Confidence Interval = .078, .093). The standardized factor loadings, item means, standard deviations, and item correlations are reported in Table 1.

TABLE 1 HERE

MacDonald’s Omega indicated good internal reliability for the PECK-SF at time 1 ($\omega = .89$) and at time 2 ($\omega = .93$), and at time 1 for both male ($\omega = .89$) and female ($\omega = .86$) participants.

Invariance testing

Regarding invariance across time, the first configural (unconstrained model) was calculated to test whether the model was a good fit and factor loadings were similar across both time conditions. This model was an acceptable fit to the data: χ^2 (df = 378) = 16,162.68, $p < .001$; CFI = .965; RMSEA = .041 (90% Confidence Interval = .037, .044). The next model assessed for weak (metric) invariance to test whether the factor loadings are statistically equivalent across conditions. If the results support weak invariance, we can infer that the factor has the same meaning across conditions (Bowen & Masa, 2015; Steinmetz, 2013), in this case over time, and for both males and female participants. The final model

PECK-SF: Tests of invariance across gender and over time.

assessed if strong (scalar) invariance was supported and involved testing whether the thresholds are equivalent. If strong invariance is found, factor means, variances, and covariances can be compared as we can infer that our observed scores provide an accurate reflection of our construct (Millsap,1998). Regarding invariance of the PECK over time, changes in RMSEA and CFI across these three models were within acceptable limits, supporting strong invariance of the scale over time (see Table 2). Regarding invariance across gender, the initial configural model could not be calculated as no female participants reported frequent experiences of item 8 (*Other kids tell people to hit me*). Therefore, this item was excluded from the model. Once this item was removed, the configural model was an acceptable fit to the data; χ^2 (df = 156) = 5,787.75, $p < .001$; CFI = .947; RMSEA = .081 (90% Confidence Interval = .072, .089). The results of the invariance tests are shown in Table 2. The changes in the fit indices were within acceptable limits across the three models and support strong invariance of the scale across gender.

TABLE 2 HERE

Discussion

This study is the first to test measurement invariance of the PECK-SF (Prinz et al., 2020) across both gender and time. Confirmatory Factor Analysis supported the one-factor structure of the PECK-SF reported by Prinz et al. (2020) and supports the use of the PECK-SF as a brief, unidimensional measure of the frequency of peer-victimization experiences. The internal reliability of the scale was good at both time points, and strong invariance across both gender and time was evidenced. These data suggest the PECK-SF provides an accurate reflection of peer-victimisation experiences and support the use of the PECK-SF when testing for gender differences, or conducting longitudinal studies, in peer-victimization experiences in early adolescence.

PECK-SF: Tests of invariance across gender and over time.

It is notable that, in our study, the PECK-SF had to be modified to produce model estimates. Specifically, initial invariance testing could not be conducted due to the low number of responses affirming more frequent (“Most days” and “Everyday”) victimization experiences. Prinz et al. (2019) reported the same issue and their solution, collapsing these categories into a single category, was also effective in the current report. The need to collapse the most frequent points on the scale is not unique to the PECK-SF, and has been reported with other measures (e.g., Due et al., 2005). Evidence suggests that those who experience very frequent peer-victimization report poorer wellbeing and different coping strategies to those who report less frequent experiences (Smith et al., 2001; Solberg & Olweus, 2003). Future research may benefit from further examination of the measurement of very frequent (chronic) experiences of peer-victimization.

It was also the case that no females in our sample reported frequent experiences (either “Most days” or “Everyday”) in response to the item “*Other kids tell people to hit me*”. This led to the omission of the item from the invariance testing relating to gender. Strong invariance across female and male participants was supported. One possible implication of this is that the item should be excluded from future use of the PECK-SF. However, we would caution against such a recommendation based on only our study. Rather, our findings should serve to emphasize that a larger sample than that recruited here may be required to exhaustively investigate the invariance of the PECK-SF in relation to gender.

While strong invariance over time was supported in the current study, the time frame was relatively short (two months). As peer-victimization can be a chronic and frequent experience for many adolescents, invariance testing over a longer period of time would be beneficial. The PECK-SF, however, is a relatively new scale, and further replication of the psychometric testing is required to lend further support to our findings. Given the importance of the development and evaluation of anti-bullying interventions, assessing the scale for

PECK-SF: Tests of invariance across gender and over time.

invariance before and after an intervention may also be important since scores changes following an intervention may reflect changes in how participants respond to the scale (e.g. if the intervention leads them to re-evaluate when constitutes aggression) rather than changes in experiences. However, the PECK-SF is a behavioral measure that requires participants to simply report the frequency of experiencing each behavior. As such, it may be less susceptible to any changes in participants understanding of the behavior due to an intervention. However, future testing the PECK-SF before and after interventions, across age groups, and cross-culturally are all valuable next steps in the development of the scale and for its application across as broad a research landscape as possible.

The PECK-SF provides a short measure of the frequency of peer-victimization, capturing different domains of victimization (physical, cultural, relational-verbal, and cyber). Confirmatory factor analysis supported the one-factor structure of the PECK-SF, and the scale was found to be reliable for male and female participants and at two-time points. Strong invariance was also supported, suggesting that the PECK-SF provides an accurate reflection of peer-victimisation and enabling researchers and clinicians to compare scores on the PECK-SF across time, and across gender. Therefore, as a short measure, the PECK-SF can provide a quick, psychometrically sound measure of the frequency of peer-victimization experiences.

PECK-SF: Tests of invariance across gender and over time.

References

- Bowen, N. K., & Masa, R. D. (2015). Conducting measurement invariance tests with ordinal data: A guide for social work researchers. *Journal of the Society for Social Work and Research, 6*(2), 229-249. <https://doi.org/10.1086/681607>
- 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
- Coertjens, L., Donche, V., De Maeyer, S., Vanthournout, G., & Van Petegem, P. (2012). Longitudinal measurement invariance of Likert-type learning strategy scales: are we using the same ruler at each wave? *Journal of Psychoeducational Assessment, 30*(6), 577-587. <https://doi.org/10.1177/0734282912438844>
- Curran, P. J., West, S. G., & Finch, J. F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological methods, 1*(1), 16-29.
- Due, P., Holstein, B. E., Lynch, J., Diderichsen, F., Gabhain, S. N., Scheidt, P., Currie, C., et al. (2005). Bullying and symptoms among school-aged children: International comparative cross sectional study in 28 countries. *European Journal of Public Health, 15*(2), 128-132. <https://doi.org/10.1093/eurpub/cki105>
- Evans, C. B., Smokowski, P. R., & Cotter, K. L. (2014). Cumulative bullying victimisation: An investigation of the dose–response relationship between victimisation and the associated mental health outcomes, social supports, and school experiences of rural adolescents. *Children and Youth Services Review, 44*, 256-264. <https://doi.org/10.1016/j.childyouth.2014.06.021>

PECK-SF: Tests of invariance across gender and over time.

- Gaffney, H., Ttofi, M. M., & Farrington, D. P. (2021). What works in anti-bullying programs? Analysis of effective intervention components. *Journal of School Psychology, 85*, 37-56. <https://doi.org/10.1016/j.jsp.2020.12.002>
- Gardella, J. H., Fisher, B. W., & Teurbe-Tolon, A. R. (2017). A systematic review and meta-analysis of cyber-victimisation and educational outcomes for adolescents. *Review of Educational Research, 87*(2), 283-308. <https://doi.org/10.3102/0034654316689136>
- Hayes, A. F., & Coutts, J. J. (2020). Use omega rather than Cronbach's alpha for estimating reliability. But.... *Communication Methods and Measures, 14*(1), 1-24.
<https://doi.org/10.1080/19312458.2020.1718629>
- Hunt, C., Peters, L., & Rapee, R. M. (2012). Development of a measure of the experience of being bullied in youth. *Psychological Assessment, 24*(1), 156-165.
<https://doi.org/10.1037/a0025178>
- Hunter, S. C., Boyle, J. M. E., & Warden, D. (2004). Help seeking amongst child and adolescent victims of peer-aggression and bullying: The influence of school-stage, gender, victimisation, appraisal, and emotion. *British Journal of Educational Psychology, 74*(3), 375-390. <https://doi.org/10.1348/0007099041552378>
- Hunter, S. C., Noret, N., & Boyle, J. M. (2021). Measurement issues relevant to questionnaire data. In P.K. Smith & J. O'Higgins Norman (Eds.). *Wiley-Blackwell Handbook of Bullying*. Wiley-Blackwell (pp179-195).
- Juvonen, J., & Graham, S. (2014). Bullying in schools: The power of bullies and the plight of victims. *Annual Review of Psychology, 65*, 159-185. <https://doi.org/10.1146/annurev-psych-010213-115030>
- Kennedy, R. S. (2019). Bullying trends in the United States: a meta-regression. *Trauma, Violence, & Abuse. <https://doi.org/10.1177/1524838019888555>*

PECK-SF: Tests of invariance across gender and over time.

Millsap, R. E. (1998). Group differences in regression intercept: Implication for factorial invariance. *Multivariate Behavioral Research*, 33, 403-424.

doi:10.1207/s15327906mbr3303_5

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>

Noret, N., Hunter, S. C., & Rasmussen, S. (in press). The role of cognitive appraisals in the relationship between peer-victimization and depressive symptomatology in adolescents:

A longitudinal study. *School Mental Health*. <https://doi.org/10.1007/s12310-021-09414-0>

Olweus, D. (1993). *Bullying at School*. Oxford: Blackwell Publishing.

Olweus, D., Limber, S. P., & Breivik, K. (2019). Addressing specific forms of bullying: A large-scale evaluation of the Olweus bullying prevention program. *International Journal of Bullying Prevention*, 1(1), 70-84. <https://doi.org/10.1007/s42380-019-00009-7>

Prinz, K., Costa, D., Chervonsky, E., & Hunt, C. (2019). Construction and psychometric properties of a short form of the Personal Experiences Checklist. *Journal of Interpersonal Violence*, <https://doi.org/10.1177/0886260519875559>.

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>

Reijntjes, A., Kamphuis, J. H., Prinzie, P., & Telch, M. J. (2010). Peer victimisation and internalising problems in children: A meta-analysis of longitudinal studies. *Child Abuse & Neglect*, 34(4), 244-252. <https://doi.org/10.1016/j.chiabu.2009.07.009>

PECK-SF: Tests of invariance across gender and over time.

- Sass, D. A., Schmitt, T. A., & Marsh, H. W. (2014). Evaluating model fit with ordered categorical data within a measurement invariance framework: A comparison of estimators. *Structural Equation Modeling: A Multidisciplinary Journal*, 21(2), 167-180. <https://doi.org/10.1080/10705511.2014.882658>
- Smith, P. K., López-Castro, L., Robinson, S., & Görzig, A. (2019). Consistency of gender differences in bullying in cross-cultural surveys. *Aggression and violent behavior*, 45, 33-40. <https://doi.org/10.1016/j.avb.2018.04.006>
- Smith, P.K., Shu, S. & Madsen, K. (2001). Characteristics of victims of school bullying: Developmental changes in coping strategies and skills. In J. Juvonen & S. Graham (Eds). *Peer-harassment in school: The plight of the vulnerable and victimised* (pp332-351). London: Guilford Press.
- Söderberg, P., & Björkqvist, K. (2020). Victimization from Peer Aggression and/or Bullying: Prevalence, Overlap, and Psychosocial Characteristics. *Journal of Aggression, Maltreatment & Trauma*, 29(2), 131-147. doi:10.1080/10926771.2019.1570410
- Steinmetz, H. (2013). Analyzing observed composite differences across groups: Is partial measurement invariance enough? *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences*, 9, 1-12.
- Trizano-Hermosilla, I., & Alvarado, J. M. (2016). Best alternatives to Cronbach's alpha reliability in realistic conditions: congeneric and asymmetrical measurements. *Frontiers in psychology*, 7, 769. <https://doi.org/10.3389/fpsyg.2016.00769>
- Valdebenito, S., Ttofi, M. M., Eisner, M., & Gaffney, H. (2017). Weapon carrying in and out of school among pure bullies, pure victims and bully-victims: A systematic review and meta-analysis of cross-sectional and longitudinal studies. *Aggression and Violent Behavior*, 33, 62-77. <https://doi.org/10.1016/j.avb.2017.01.004>

PECK-SF: Tests of invariance across gender and over time.

Vessey, J., Strout, T. D., DiFazio, R. L., & Walker, A. (2014). Measuring the youth bullying experience: A systematic review of the psychometric properties of available instruments. *Journal of School Health, 84*(12), 819-843.

<https://doi.org/10.1111/josh.12210>

Volk, A. A., Veenstra, R., & Espelage, D. L. (2017). So you want to study bullying? Recommendations to enhance the validity, transparency, and compatibility of bullying research. *Aggression and Violent Behavior, 36*, 34-43.

<https://doi.org/10.1016/j.avb.2017.07.003>

Zych, I., Ttofi, M. M., Llorent, V. J., Farrington, D. P., Ribeaud, D., & Eisner, M. P. (2020).

A longitudinal study on stability and transitions among bullying roles. *Child Development, 91*(2), 527-545. <https://doi.org/10.1111/cdev.13195>

PECK-SF: Tests of invariance across gender and over time.

Table 1: *Factor Loadings from Confirmatory Factor Analysis, Item Correlations, and Descriptive Statistics for all Items*

	Factor Loadings	Mean (SD)	Item Correlations												
			2	3	4	5	6	7	8	9	10	11	12	13	14
1. The other kids ignore me on purpose.	0.74	1.77 (0.71)	.23***	.49***	.34***	.54***	.27***	.42***	.36***	.23***	.58***	.41***	.38***	.46***	.37***
2. The other kids make fun of my language.	0.51	1.21 (0.38)	-	.25***	.22***	.26***	.38***	.20***	.32***	.25***	.25***	.25***	.20***	.28***	.23***
3. The other kids tease me about things that aren't true.	0.71	1.88 (0.94)	-	-	.33***	.38***	.27***	.43***	.32***	.21***	.56***	.41***	.43***	.44***	.32***
4. Other kids threaten me over the phone.	0.79	1.18 (0.30)	-	-	-	.46***	.43***	.55***	.48***	.48***	.38***	.32***	.54***	.41***	.29***
5. Other kids tell people not to hang around with me.	0.82	1.33 (0.48)	-	-	-	-	.42***	.43***	.46***	.32***	.57***	.40***	.43***	.51***	.26***
6. Other kids won't talk to me because of where I'm from.	0.81	1.07 (0.14)	-	-	-	-	-	.31***	.52***	.40***	.27***	.32***	.28***	.39***	.29***
7. Other kids say nasty things to me by texting.	0.81	1.38 (1.38)	-	-	-	-	-	-	.40***	.41***	.53***	.29***	.65***	.45***	.25***
8. Other kids tell people to hit me	0.85	1.16 (0.26)	-	-	-	-	-	-	-	.42***	.35***	.49***	.43***	.56***	.57***
9. Other kids send me nasty e-mails.	0.73	1.09 (0.17)	-	-	-	-	-	-	-	-	.24***	.26***	.40***	.38***	.30***
10. Other kids say mean things behind my back.	0.84	1.95 (1.05)	-	-	-	-	-	-	-	-	-	.43***	.54***	.49***	.27***
11. Other kids shove me.	0.74	1.45 (0.65)	-	-	-	-	-	-	-	-	-	-	.36***	.47***	.56***
12. Other kids say nasty things about me on online.	0.84	1.34 (0.52)	-	-	-	-	-	-	-	-	-	-	-	.51***	.29***
13. Other kids tell people to make fun of me.	0.83	1.27 (0.41)	-	-	-	-	-	-	-	-	-	-	-	-	.43***
14. Other kids hit me.	0.71	1.26 (0.44)	-	-	-	-	-	-	-	-	-	-	-	-	-

***p<.001

PECK-SF: Tests of invariance across gender and over time.

Table 2: Tests of invariance for the PECK-SF

Model	CFI		Δ CFI		RMSEA (90% CI)		Δ RMSEA	
			(vs. preceding model)				(vs. preceding model)	
	Time	Gender	Time	Gender	Time	Gender	Time	Gender
Model 1 (Unconstrained)	.965	.947			.041 (.037, .044)	.081 (.072, .089)		
Model 2 (Weak invariance)	.971	.936	+0.006	-0.011	.037 (.033, .040)	.085 (.077, .093)	-0.004	-0.004
Model 3 (Strong invariance)	.970	.925	-0.001	-0.011	.035 (.032, .039)	.082 (.075, .089)	-0.002	-0.003