The Impact of Validity Screening on Associations Between Self-Reports of Bullying Victimization and Student Outcomes

Yuane Jia1, Timothy R. Konold1, Dewey Cornell1 & Francis Huang2

1 Curry School of Education, University of Virginia
2 College of Education, University of Missouri–Columbia

Abstract

Self-report surveys are widely used to measure adolescent risk behavior and academic adjustment, with results having an impact on national policy, assessment of school quality, and evaluation of school interventions. However, data obtained from self-reports can be distorted when adolescents intentionally provide inaccurate or careless responses. The current study illustrates the problem of invalid respondents in a sample (N = 52,012) from 323 high schools that responded to a statewide assessment of school climate. Two approaches for identifying invalid respondents were applied, and contrasts between the valid and invalid responses revealed differences in means, prevalence rates of student adjustment, and associations among reports of bullying victimization and student adjustment outcomes. The results lend additional support for the need to screen for invalid responders in adolescent samples.

Background

• Failure to screen samples for invalid responders has been found to lead to both exaggerated prevalence rates (Cornell et al., 2012; Furlong, Sharkey, Bates, & Smith, 2008) and erroneous conclusions regarding associations between student conditions (e.g., adoption, disability status) and adjustment (e.g. drinking, fighting, low self-esteem, low school engagement; Fan et al., 2002-2006).
• No studies of bullying have examined the use of survey completion time to identify invalid responders. The combination of validity screening items and assessment of survey completion time might provide a more effective way to identify survey data that should be omitted from analyses.
• Two ways of identifying invalid responders were used: built-in validity screening items and survey completion time.

Methods

Sample and Measures

• Data were from a statewide sample completing the Authoritative School Climate (ASC) Survey (Cornell et al., 2014). 52,012 students from 323 public high schools responded through a secure online Qualtrics platform. The validity screening items flagged 3,579 students (6.88% of the sample). And 649 students were identified as completing the survey too rapidly (less than 6.1 minutes). In total, 3,985 responders were classified in the invalid group (7.66% of the total sample), leaving 48,027 in the valid group.
• Bullying victimization, student engagement, and GPA were obtained from the ASC survey and risk behaviors were measured using items from the Youth Risk Behavior Survey (for more details see Cornell et al., 2014).
• Validity screen items: “I am telling the truth on this survey”; How many of the questions on this survey did you answer truthfully?” Survey completion time was recorded by the Qualtrics platform.

Results

Table 4. Logistic Regression for Predicting Dichotomous Outcomes of Student Adjustment.

<table>
<thead>
<tr>
<th>Step</th>
<th>OR</th>
<th>CI</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Results of the logistic regression analysis indicated that the combination of validity screening items and survey completion time was effective in identifying invalid responders. The results showed significant differences in means, prevalence rates, and correlations among the valid and invalid groups. The findings also highlighted the importance of considering both validity screening items and survey completion time in the identification of invalid respondents.

Conclusions and implications

• Inclusion of the invalid responders in the total sample inflated the prevalence of all reported risk behaviors with the exception of suicidal thoughts, and deflated student reports of GPA, school engagement, and depression. More inflation was found with binary outcomes.
• All adjusted mean differences between valid/invalid groups across student outcomes were statistically significant (p<.001), with meaningful effect sizes.
• Bullying-adjustment associations were moderated by valid vs. invalid group membership across all investigated outcomes, except for depression. The magnitudes of the differences between the valid and invalid groups were dependent on the level of reported bullying victimization (see figure 1).
• Future research employing self-report surveys with adolescents should consider incorporating some form of validity screening to avoid both inflated prevalence rates and correlations.