Brief/Psychometric Report

Hammill Institute on Disabilities

Assessment for Effective Intervention
2022, Vol. 47(2) 112–117

DOI: 10.1177/1534508420959439
aei.sagepub.com
SAGE

Anonymous Versus Self-Identified Response Formats for School Mental Health Screening

Rhea Wagle, MEd[1], Erin Dowdy, PhD[1], Michael J. Furlong, PhD[1], Karen Nylund-Gibson, PhD[1], Delwin Carter, MEd[1], and Tameisha Hinton, MEd[1]

Abstract

Schools are an essential setting for mental health supports and services for students. To support student well-being, schools engage in universal mental health screening to identify students in need of support and to provide surveillance data for district-wide or state-wide policy changes. Mental health data have been collected via anonymous and self-identified response formats depending on the purpose of the screening (i.e., surveillance and screening, respectively). However, most surveys do not provide psychometric evidence for use in both types of response formats. The current study examined whether responses to the Social Emotional Health Survey–Secondary (SEHS-S), a school mental health survey, are comparable when administered using anonymous versus self-identified response formats. The study participants were from one high school and completed the SEHS-S using self-identified (n = 1,700) and anonymous (n = 1,667) formats. Full measurement invariance was found across the two response formats. Both substantial and minimal latent mean differences were detected. Implications for the use and interpretation of the SEHS-S for schoolwide mental health are discussed.

Keywords

universal screening, school-based, anonymous, identified, Social Emotional Health Survey–Secondary

Universal screening is often the first step in identifying, preventing, and treating mental health problems (Levitt et al., 2007), and the number of schools participating in universal mental health screening has increased over the past decade (Bruhn et al., 2014). Universal screening assesses an entire population, often via student self-report, whereby students either anonymously or via identifiable means report on their mental health functioning. There are two main approaches for school-based universal mental health screening, aligned with principles of public health and population-based assessment (Doll & Cummings, 2008). One approach is to gather information on students' mental health to direct prevention and treatment services for specific students. When schools engage in screening to identify specific at-risk students, the responses need to be self-identifiable; that is, when providing assent and before responding, the student is aware that their answers will be known, confidentially, to responsible school staff. The other approach gathers information about an entire population for surveillance purposes, often to aid in developing school-wide or district-wide prevention and intervention services in addition to policy changes. Schools interested in surveillance data can employ an anonymous survey format. Regardless of the primary purpose, it is critical to understand better how the response format might affect the quality of student responses. This report examines anonymous and self-identified survey formats within the context of universal school-based mental health screening.

Past research has identified survey response differences depending on whether they are collected using self-identified or anonymous formats. This research shows that participants tend to disclose higher degrees of stigmatizing or sensitive information (i.e., mental health symptoms) on anonymous rather than identifiable surveys (e.g., Beebe et al., 2006). Social desirability bias, or the tendency for participants to endorse more socially desirable responses, can affect respondents' truthfulness when answering surveys in a non-anonymous format. Several studies found that participants self-reported lower social anxiety and social desirability when they were anonymous than when they were non-anonymous (e.g., Joinson, 1999).

[1]University of California, Santa Barbara, USA

Corresponding Author:
Rhea Wagle, Department of Counseling Clinical and School Psychology, University of California, Santa Barbara, Bldg. 275, Room 4100, Santa Barbara, CA 93106-9490, USA.
Email: rheawagle@ucsb.edu

Furthermore, individuals may provide different responses to surveys based on degree of anonymity. Chandler and colleagues (2020) found that fraudulent or careless survey responses, which tended to be positively skewed, can lead to data-quality problems, such as false-positive between-group differences. With screening efforts in schools having important implications, there is a need to understand the consistency of students' responses in the school-based mental health screening context. To meaningfully interpret survey results, it is critical to understand whether measures assess the same constructs when employing anonymous versus self-identifiable survey formats.

There are several school-based universal mental health screening measures (e.g., Bates & McKay Boren, 2020). One such measure, the Social Emotional Health Survey–Secondary (SEHS-S; Furlong et al., 2014), assesses students' social emotional strengths. Utilizing a strength-based survey rather than a traditional deficit-focused assessment provides a holistic understanding of students' well-being, consistent with contemporary views of mental health (Suldo & Shaffer, 2008). Previous studies have produced evidence supporting reliability and validity of the SEHS-S across sociocultural and gender groups, for use in screening to identify specific students for intervention (e.g., You et al., 2015) and for schoolwide and district-wide surveillance (California Healthy Kids Survey, https://calschls.org/). Nevertheless, to date, no study has examined how, or if, students' SEHS-S responses are comparable across survey formats. To better understand the results of surveillance and screening applications of the SEHS-S, it is crucial to examine survey format comparability, so school personnel can draw appropriate comparisons when anonymous versus self-identified survey formats are employed. The current exploratory study sought to contribute to school-based universal mental health screening research and practice by answering the following: (a) Does the SEHS-S provide psychometrically comparable (i.e., invariant) information when used for both surveillance (anonymous) and specific-student (self-identified) purposes? That is, is there evidence of measurement invariance (MI) across these groups?; (b) If MI is found, are students' responses comparable based on whether they are anonymous or self-identified?

Method

Participants

Students in Grades 9 to 12 from one high school in Central California completed surveys during the 2017–2018 school year. The self-identified group included 1,700 (81% of enrolled) students who were surveyed in fall 2017. These students were evenly distributed across grades (ninth grade = 27.3%, 10th grade = 26.1%, 11th grade = 13%, 12th grade = 23.6%) and gender (female = 52.3%, male = 45.9%, other = 1.8%). The ethnic makeup was as follows: Latinx = 44.9%, White = 40.2%, multiracial = 8.7%, Asian = 2.9%, Black = 1.4%, Native American = 1.4%, and Pacific Islander = 0.5%. Drawing from the students at the same high school, 1,667 (79% of enrolled) students completed the same survey items using an anonymous survey administration procedure in spring 2018. Because the anonymous response format did not include a unique identifier, it was not possible to determine how many students completed both formats. However, given the known total school enrollment, the overlap was substantial (possible range: 75%–99%). No significant differences across the self-identified and anonymous samples were found in the proportion of students for grade level, $t = 0.38$, $p = .70$, or gender identification, $t = -1.08$, $p = .28$. Differences in ethnic identification could not be tested due to differing response options available for each sample.

Measures

The SEHS-S (Furlong et al., 2020, 2014) is a 36-item self-report measure used to assess the social emotional strengths of secondary students (Grades 6–12). Previous factor analytic studies have supported a higher order-factor structure with 36 items loading onto 12 subscales, which subsequently load onto four second-order traits (e.g., domains) of *belief in self* (*self-awareness, persistence, self-efficacy*), *belief in others* (*school support, family coherence, peer support*), *emotional competence* (*empathy, self-control, behavioral self-control*), and *engaged living* (*gratitude, zest*, and *optimism*). This model is called the *four-factor correlated model*. Another model that studies have evaluated is the *full model*, where the four second-order traits then load onto a higher order construct called *covitality*. Both models have evidence of validity and reliability through confirmatory factor analyses (CFA) and MI analyses (e.g., You et al., 2015). The current study examined both models.

Procedure

Students at one comprehensive high school completed the surveys during the 2017–2018 school year. Students completed the survey using a self-identified format as part of a federally funded study investigating the psychometric properties of the SEHS-S (fall 2017). The school was interested in acquiring information for school staff to follow up with students through traditional counseling and support services. Students at the same high school later (spring 2018) completed the SEHS-S using an anonymous survey format and as part of a state-wide administration of the California Healthy Kids Survey through WestEd.

Self-identified. The school district approved the use of passive parental consent and student assent. Consent and assent

Table 1. Model Fit Statistics of Factor Models With Differing Levels of Measurement Invariance.

Models	CFI	RMSEA	SRMR	χ^2	*df*	$\Delta\chi^2$	Δdf	ΔCFI	ΔRMSEA
CFA									
Both	.964	.038	.04	3420.67	576				
Anonymous	.958	.048	.04	2821.23	576				
Self-identified	.953	.037	.04	1965.60	576				
MI Level 1									
Configural	.968	.039	.03	3730.23	1,056				
Metric	.967	.039	.03	3803.47	1,080	73.00***	24	.001	.000
Scalar	.961	.041	.03	4305.69	1,104	502.22***	15	.006	.002
MI Level 2									
Configural	.950	.045	.04	5279.23	1,176				
Metric	.950	.045	.05	5361.24	1,200	81.91***	24	.000	.000
Scalar	.948	.046	.05	5499.94	1,208	138.70***	8	.002	.001

Note. CFI = comparative fit index; RMSEA = root mean square of approximation; SRMR = standardized root mean square residual; CFA = confirmatory factor analysis; MI = measurement invariance.
***$p < .001$.

procedures were available in Spanish and English. Three parents declined consent, and 150 students declined assent. All students with parental consent and assent completed the survey via computers in the school computer lab or on tablets in the classroom. Students had the option of using a toggle function to view items in Spanish and English. Classroom teachers and researchers proctored the administration using a standardized script that explained the nature of the survey to all students. The online survey format explained the survey purpose and asked students to enter their unique school ID number. Students were told that school staff would be able to review their responses consistent with the purpose of the survey:

> Who will see my answers? The school staff will not share your answers with anyone. If the school staff think you might benefit from extra support, they will meet with you so that they can figure out what will be most helpful.

Anonymous. Consistent with school procedures to complete the California Healthy Kids Survey (CHKS; see https://calschls.org/survey-administration/parental-consent/), the school district allowed for passive parental consent. Students whose parents did not opt-out participated in the survey using computers in a computer lab or using personal tablets in the classroom. Teachers proctored the survey administration and utilized a standardized script to explain the purpose of the survey. Specifically, students were informed

> You do not have to answer these questions, but your answers will be very helpful in improving school and health programs. The survey is anonymous and confidential. No one will be able to connect you with your answers. Your answers are private.

Teachers were available to answer any questions.

Statistical Analyses

Data quality checks. Wested's 10-item case reject index (response inconsistency, fictitious drug use, excessive alcohol and other drug (AOD) use, and response dishonesty; https://calschls.org/docs/validity.pdf) identified 30 students for exclusion from the Anonymous sample. Normality assumptions were tested and met by analyzing descriptive statistics of all survey items before investigating between-group MI (see Table 1).

MI analyses. MI tests across both groups were completed to evaluate whether the SEHS-S items relate to the factors in the same way based on anonymity of responses in Mplus (Muthén & Muthén, 1998-2017). The CFA model fit was evaluated using recommendations from Hu and Bentler (1999) and Browne and Cudeck (1989). More specifically, the model fit was found by identifying comparative fit index (CFI) values above .95 to indicate good fit and values above .90 to indicate adequate fit (Hu & Bentler, 1999). Consistent with procedures outlined in Browne and Cudeck (1989), root mean square of approximation (RMSEA) and standardized root mean square residual (SRMR) values less than .05 suggest good fit and values up to .08 suggest reasonable fit.

A series of CFA were fit. First, we fit a CFA model where each group's parameters were estimated independently, though simultaneous (e.g., multiple groups). This step established configural invariance. Next, item loadings were constrained to be equal across both groups, while all other model parameters were freely estimated to establish metric invariance. Finally, scalar invariance was tested by fixing item intercepts and loadings to be equal across both groups. Nested models were compared using the chi-square difference test ($\Delta\chi^2$; Chen, 2007) and analyzing the change in

CFI (ΔCFI) and change in RMSEA (ΔRMSEA), such that values of $\Delta CFI \leq .01$ and $\Delta RMSEA \leq .015$ supported MI (Chen, 2007). If MI was found, latent mean comparisons were made to understand the difference in factor means.

Results and Discussion

Confirmatory Factor Analysis

A series of CFAs were conducted across both groups of students (i.e., anonymous and self-identified). First, a CFA analyzed model fit for the full model, inclusive of the *covitality* construct. The CFA for the self-identified group did not converge due to high collinearity of the higher order factors. This result might be due to the influence of social desirability bias or other factors from producing self-identified responses. Additional research is required to understand further why this model did not converge with the self-identified sample. Thus, the analyses continued with the four-factor correlated model. The CFAs of the anonymous group, self-identified group, and both together found a good model fit for the four-factor correlated model (see Table 1). Past studies found similar factor structures and model fit information when analyzing the SEHS-S with various populations (e.g., You et al., 2015).

Measurement of Invariance

Three levels of invariance were tested: configural, metric, and scalar invariance. In addition, due to the higher order nature of the factor structure (i.e., 36 items loading onto 12 subscales, which then load onto four factors), each invariance analysis was tested at the first level (i.e., 36 items loading onto 12 subscales) and the second level (i.e., 12 subscales loading onto four factors). All three levels of invariance found good model fit for both the first-level model and the second-level model with ($\Delta CFI < .01$, $\Delta RMSEA < .01$) between levels indicating that the constraints did not lead to a meaningful change hence reaching full MI (see Table 1)

Thus, full MI was assumed at both factor levels across the anonymous and self-identified groups for the four-factor correlated model. These results suggest that the same construct is measured across both groups. This finding provided evidence supporting the use of SEHS-S to screen for and provide services to specific students (self-identified format) and for surveillance purposes that support district-wide or state-wide policy initiatives (anonymous format). These findings echo similar studies that have reported MI for the SEHS-S across multiple groups such as gender, age, and race/ethnicity (e.g., Furlong et al., 2014; You at al., 2015).

Latent Mean Differences

With MI established, latent means were compared for anonymous and self-identified formats. With the anonymous group serving as the reference group, the latent means for the self-identified group were freely estimated (Byrne, 2012). There were mean differences for six of the 12 subscales and three of the four domains. Students in the self-identified group reported higher *self-awareness* (ß = .29, $p < .001$), *school support* (ß = .20, $p < .001$), *family coherence* (ß = .18, $p < .001$), *self-control* (ß = .13, $p = .015$), *optimism* (ß = .14, $p = .003$), and *gratitude* (ß = .15, $p < .001$). There were no significant differences for *persistence, self-efficacy, peer-support*, e*mpathy, behavioral self-control*, and *zest*.

At the second level (i.e., 12 subscales loading onto four factors), the analyses showed significant mean differences for *belief in self* (ß = .11, $p = .023$), *belief in others* (ß = .25, $p < .001$), and *emotional competence* (ß = .11, $p = .021$). Students in the self-identified group reported higher *belief in self, belief in others*, and *emotional competence* than those in the anonymous group. There was no significant difference in *engaged living*.

Some subscales and factors showed significant latent mean differences with variable effect sizes (between .11 and .29), indicating small to large latent mean differences. In addition, students in the self-identified group tended to self-report higher scores in certain areas, consistent with previous studies (Beebe et al., 2006; Gordon, 1987). Among the six subscales and three factors with significant latent mean differences, the largest effect sizes were within *self-awareness, school support*, and *belief in others*, indicating that these areas were likely the most influenced by lack of anonymity, possibly including a social desirability bias. *School support*, also a component of the *belief in others* factor, is an expected area of difference because students might inflate opinions about school when school personnel can identify their responses. The differences in *self-awareness* were less expected, though they may be attributable to social desirability bias and self-image management.

The other significant latent mean differences were negligible with small effect sizes. Six other subscales and one factor had nonsignificant latent mean differences. These results suggest that not all information on student mental health is lost, changed, or rendered ambiguous by asking students to provide self-identifying information. Moreover, considering the benefits of self-identifying student responses (e.g., providing services to specific students in need), there is a compelling rationale for schools to use self-identification when implementing universal screening. Doing so would allow schools to provide treatment and care for their students while also gaining critical surveillance information for policy planning; that is, the use of a self-identification survey format could service both screening and surveillance purposes.

The present study focused on response differences between anonymous and self-identified groups. Future studies could examine the effects of other forms of response bias. For example, future studies could examine anonymity's effects on response honesty. Future research could also examine consent or assent differences. For example, does a self-identified survey format decrease parental consent and student assent? Finally, research is needed to determine whether these findings replicate in other diverse populations.

A limitation of the present study is that the analyses employed opportunity, not preplanned samples. It could not be determined how many students completed *both* the anonymous and self-identified survey formats. Yet, an ethically defensible and approved survey to conduct such a preplanned study that would link anonymous and self-identified responses would be impractical—the use of self-identification for screening purposes must include a statement asking students to enter an ID useable by school staff, while also informing them school personnel will be able to see and review their responses. Consequently, a pre-planned survey format study that included any form of deception or deceptive response matching would be unethical. The goal of this study was to evaluate student responses under the necessary condition that they specifically knew that school staff would be able to see their responses.

Furthermore, a pre-planned study would forgo other limitations such as the lag in time between data collection for each sample, which may have contributed to response differences because some students participated in service-as-usual social emotional supports provided by the school. While there is no information about the types of social emotional supports available to individual students at this school, past research shows the SEHS-S scores are trait-like with high 1-year stability coefficients (Furlong et al., 2020), indicating that the survey responses would likely not differ significantly within this study's lag time. Finally, data were collected from a single school, providing a limitation regarding the generalizability of the findings. Analyzing data from several schools in future studies may provide increased support for the findings. Despite these limitations, the current study provided an opportunity to explore how an anonymous response format affected students' responses to universal mental health screening.

Conclusion

The current study analyzed whether the SEHS-S, a measure used to assess students' well-being, is valid when used with anonymous and self-identified survey formats. Findings indicated that the SEHS-S had full MI for anonymous and self-identified samples, supporting its use for surveillance and universal screening purposes. There were significant latent mean differences across some SEHS-S subscales and factors, with effect sizes ranging from minimal to large. The self-identified group reported higher scores. For example, students in the self-identified group reported higher *self-awareness* and *belief in self*, which included *school support.* It is important to note that students' inflated scores, mainly on items asking about school support, when their responses are self-identifiable. This could be due to social desirability bias. An alternative hypothesis is that it could be influenced by the level of trust students have with school staff. To maximize the quality and utility of self-identified surveys, schools should consider collecting schoolwide screening data through a broader multi-tiered support structure to foster trust and positive student-school staff relationships.

Additional significant latent mean differences had small effect sizes and six subscales, and one factor showed nonsignificant latent mean differences. Thus, this study's main implication is that it provided preliminary evidence that universal self-identified school mental health surveys provide information that is comparable to surveys that employ anonymous formats. This finding can assuage school concerns that, when asked about their mental health experiences, students will self-disclose in a meaningful and useful manner. It is important to note that the results of this study are specific for the SEHS-S and might generalize to other strength-focused screening measures but not pathology (e.g., diagnostic depression) measures.

Support for comprehensive school mental health services is enhanced when universal screening survey procedures ask students to provide self-identifying information. School care teams can then respectfully monitor students and respond to foster their well-being. When self-identified surveys are integrated across local and state education agencies, they also provide surveillance information that informs mental health policies and legislation. Finally, research needs to replicate these findings with diverse populations with in-depth examinations of how anonymity and other response biases influence student responses.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The research reported here was supported in part by the Institute of Education Sciences, U.S. Department of Education, through Grant R305A160157 to the University of California, Santa Barbara. The opinions expressed are those of the authors and do not represent views of the Institute of Education Sciences or the U.S. Department of Education.

ORCID iD

Rhea Wagle https://orcid.org/0000-0002-5402-5882

References

Bates, M., & McKay Boren, D. (2020). *Assessing wellbeing in schools: An educators practical guide to measuring wellbeing in schools*. EdTech Books. https://edtechbooks.org/wellbeing

Beebe, T. J., Harrison, P. A., Park, E., McRae, J. A., Jr., & Evans, J. (2006). The effects of data collection mode and disclosure on adolescent reporting of health behavior. *Social Science Computer Review*, *24*(4), 476–488. https://doi.org/10.1177/0894439306288690

Browne, M. W., & Cudeck, R. (1989). Single sample cross-validation indices for covariance structures. *Multivariate Behavioral Research*, *24*(4), 445–455. https://doi.org/10.1207/s15327906mbr2404_4

Bruhn, A. L., Woods-Groves, S., & Huddle, S. (2014). A preliminary investigation of emotional and behavioral screening practices in K-12 schools. *Education and Treatment of Children*, *37*(4), 611–634. https://doi.org/10.1353/etc.2014.0039

Byrne, B. M. (2012). *Multivariate applications series.Structural equation modeling with Mplus: Basic concepts, applications, and programming*. Routledge/Taylor & Francis Group.

Chandler, J., Sisso, I., & Shapiro, D. (2020). Participant carelessness and fraud: Consequences for clinical research and potential solutions. *Journal of Abnormal Psychology*, *129*(1), 49–55. https://doi.org/10.1037/abn0000479

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, *14*(3), 464–504. https://doi.org/10.1080/10705510701301834

Doll, B., & Cummings, J. (2008). Best practices in population-based school mental health services. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology V* (pp. 1333–1347). National Association of School Psychologists.

Furlong, M. J., Nylund-Gibson, K., Dowdy, E., Wagle, R., Hinton, T., & Carter, D. (2020). *Modification and standardization of Social Emotional Health Survey-Secondary-2020 edition*. https://www.researchgate.net/publication/337261781_Modification_and_standardization_of_Social_Emotional_Health_Survey-Secondary-2020_edition

Furlong, M. J., You, S., Renshaw, T. L., Smith, D. C., & O'Malley, M. D. (2014). Preliminary development and validation of the Social and Emotional Health Survey for secondary students. *Social Indicators Research*, *117*, 1011–1032. https://doi.org/10.1007/s11205-013-0373-0

Gordon, R. A. (1987). Social desirability bias: A demonstration and technique for its reduction. *Teaching of Psychology, 14*(1), 40–42. https://doi.org/10.1207/s15328023top1401_11

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*, 1–55. https://doi.org/10.1080/10705519909540011 8

Joinson, A. (1999). Social desirability, anonymity, and Internet-based questionnaires. *Behavior Research Methods, Instruments, & Computers*, *31*(3), 433–438. https://doi.org/10.3758/BF03200723

Levitt, J. M., Saka, N., Romanelli, L. H., & Hoagwood, K. (2007). Early identification of mental health problems in schools: The status of instrumentation. *Journal of School Psychology*, *45*(2), 163–191. https://doi.org/10.1016/j.jsp.2006.11.005

Muthén, L. K., & Muthén, B. O. (1998-2017). *Mplus user's guide* (8th ed.).

Suldo, S. M., & Shaffer, E. J. (2008). Looking beyond psychopathology: The dual-factor model of mental health in youth. *School Psychology Review*, *37*(1), 52–68.

You, S., Furlong, M. J., Felix, E., & O'Malley, M. D. (2015). Validation of the Social and Emotional Health Survey for five sociocultural groups: Multigroup invariance and latent mean analyses. *Psychology in the Schools*, *52*, 349–362. https://doi.org/10.1002/pits.21828/abstract