

Review of **School Situation Survey**, by Theodore Coladarci¹

The School Situation Survey (SSS) is a 34-item instrument designed to assess “school-related student stress” in grades 4 through 12. Comprising seven scales, the SSS provides separate scores for four “sources” of stress: Teacher Interactions (“students’ perceptions of their teachers’ attitudes toward them”), Academic Stress (“situations that relate to academic performance or achievement”), Peer Interactions (“students’ social interactions or their perceptions of their classmates’ feelings toward them”), and Academic Self-Concept (“students’ feelings of self-worth, self-esteem, or self-concept relevant to perceived ability”). The remaining three scales yield scores regarding “manifestations” of stress: Emotional (“feelings such as fear, shyness, and loneliness”), Behavioral (“actions, reactions, or behavior toward others, such as striking out or being hurtful or disrespectful”), and Physiological (“physical reactions or functions such as nausea, tremors, or rapid heart beat”).

Students rate each item (e.g., “I feel upset”) on a 5-point scale ranging from never to always. Answer sheets can be hand scored easily with the accompanying acetate overlay, or returned to Consulting Psychologists Press for machine scoring. In either case, a student receives seven scores; a composite score across the seven scales is not provided (nor do the authors recommend that such a score be used).

DOCUMENTATION. The SSS manual is a physically attractive document containing helpful information for SSS users and appraisers alike. However, I found the manual wanting in several respects. First, little attention is devoted to the intended uses, and possible misuses, of the SSS. For example, it would seem that different concerns would surface for practitioners than for researchers, yet the authors say little for either audience.

Second, the overall organization of the manual is awkward at times. Suggested strategies for reducing stress are presented after a section on norms, and information regarding factor analyses is alluded to in the construct validity section--where it belongs--but the reader must refer back to an earlier section on item development to see the specifics.

Finally, the authors’ treatment of fundamental aspects of instrumentation (e.g., validity) appears to be intended for readers with a technical background. Although the language of psychometrics and statistics is unavoidable in a users’ manual, it serves an important educative purpose when presented well. I fear that some of the authors’ language will have the unintended effect of distancing many readers from the information needed to make an informed appraisal of the SSS.

STANDARDIZATION SAMPLE. Norms are based on 7,036 students from grades 3 through 12 in 16 Connecticut and Rhode Island school districts. In two appendices, the authors report means, standard deviations, and ranges for interpreting scores for each of four grade-level clusters: grades 3-5, 6-8, 9, and 10-12. Within each grade level, descriptive statistics are broken down by sex.

The impressive number of students notwithstanding, these norms should be used cautiously. Schools were drawn from rural, suburban, and urban communities, but nothing is said about

¹ Coladarci, T. (1995). Review of School Situation Survey. In J. C. Conoley & J. C. Impara (eds.), *The Twelfth Mental Measurements Yearbook* (pp. 928-931), Lincoln, NE: The Buros Institute of Mental Measurements.

demographic considerations such as ethnicity and socioeconomic status, both of which would seem relevant to the problem of school-related stress. Consequently, the degree to which these norms are appropriate for any one school district remains an open question.

The authors, furthermore, do not report the number of students who were sampled in each grade. Can we assume the sixth-, seventh-, and eighth-grade students are evenly distributed within the 6-8-grade-level cluster? This ambiguity has implications for how confidently the SSS norms can be employed for a particular grade. (Curiously, the authors do not explain why they included third-grade students in the norming group for a survey designed for grades 4 through 12.)

Finally, we are told that SSS norms can be used “to provide feedback on an individual or *group* [italics added] basis” (p. 7). In fact, because the norms are based on individuals, they cannot be used to form judgments about groups of individuals (e.g., students in a particular grade or building). Consequently, the authors’ three categories for interpreting scores (low, medium, high), derived from individual-level distributions of scores, are inappropriate for interpreting group data.

RELIABILITY. Internal-consistency coefficients for the seven scales are moderate, ranging from .68 to .80 when based on the entire sample of 7,036 students; similar coefficients are reported within each of the four grade-level clusters. Importantly, a standard error of measurement (SEM) is reported for each coefficient. The SEMs, expressed on the SSS 5-point scale, range from roughly one-third to over one-half of a point (.31 to .58), depending on the scale and grade-level cluster. Combined, these data raise questions about the suitability of the SSS for forming judgments and making decisions about students, particularly at the individual level.

Test-retest reliability was determined over a 3-week period for a sample of seventh- to ninth-grade students ($n = 621$), resulting in coefficients ranging from .61 to .71. The lower value of these coefficients is not surprising insofar as affective characteristics are less stable than cognitive aptitudes or academic achievement. Users nonetheless should realize the implication: Students’ perceptions of school-related stress today, as measured by the SSS, might well differ from their perceptions in a few weeks.

We do not know how these 621 students were selected for the test-retest reliability analyses, or why a group of students was designated that cuts across two of the four grade-level clusters. More troubling, however, the authors do not explain why test-retest reliabilities are not reported for students in the remaining grade-level clusters. Can we assume, for example, that comparable coefficients would obtain for the youngest cluster? My sense is that the stress perceptions among students in grades 3 through 5 would be less stable than for older students. In any case, the stability of the SSS remains unestablished for a large segment of the target population.

VALIDITY. The authors invested considerable time and effort in the validation of the SSS.

Content validity. The content-validity question is raised by the authors this way: “To what extent do the items of [the SSS] adequately sample from the intended content domain?” (p. 14). The reader is then referred back to an earlier section on item development, which touches on the authors’ procedure for constructing an initial set of 56 items by consulting the literature and talking with groups of students, parents, educators, and specialists. In my view, this discussion falls short of establishing the universe from which these 56 items were drawn. Consequently, it is

difficult to determine how well (a) each scale samples the domain of items represented by that scale or (b) how well the seven SSS scales represent all sources and manifestations of stress.

Construct validity. According to Messick (1989), “construct validity is evaluated by investigating what qualities a test measures, that is, by determining the degree to which certain explanatory concepts or constructs account for performance on the test” (p. 16). The authors approached this task in three ways.

Factor analysis. The initial 56 items were factor analyzed on a sample of 907 students from grades 5, 7, and 9. After some revisions, the instrument was administered to a new sample of 1,111 students (also from grades 5, 7, and 9) and again factor analyzed. From this second set of analyses, 7 of the 14 obtained factors “replicated the original constructs” (pp. 11-12) that earlier had emerged from the literature and the authors’ discussions with various groups. It is these seven factors, and their 34 items, that constitute the SSS.

These factors, with the corresponding items and loadings, are clearly presented in the manual. Importantly, factor loadings show a relatively clean separation of the seven factors. And the manner in which the items cluster within each factor makes conceptual sense, as well. For example, each of the three items in the Academic Stress factor pertains to achievement situations in school. Together, these data support the authors’ claim that the SSS gets at relatively distinct aspects of students’ stress perceptions.

Surprisingly, no factor analyses were conducted on the six remaining grades for which the SSS is intended. Although I have no a priori reason to question the validity of this instrument for these six grades, their absence precludes an adequate appraisal of the construct validity of the SSS. And why did the authors not base their factor analyses on the standardization sample of 7,036 students, rather than the restricted sample of 1,111? After all, the standardization sample provides the basis for the norms and internal-consistency reliability estimates. It would seem that the authors passed up an important opportunity for cross-validation, based on a larger sample, and with all grades represented.

Correlations with the State-Trait Anxiety Inventory for Children. Any instrument, of course, should correlate with existing measures of similar constructs. Consequently, the authors correlated each of the seven SSS scales with the A-Trait scale of the State-Trait Anxiety Inventory for Children (STAIC; Spielberger, Edwards, Lushene, Montuori, & Platzek, 1973), based on the restricted sample of 1,111 students from grades 5, 7, and 9. The obtained correlations range considerably: r s = .10 to .71, with a median r of .33. In general, these correlations provide weak to moderate support for the construct validity of the SSS.

It also would be informative to know whether the SSS scales fail to correlate with measures of constructs dissimilar to school-related student stress. That is, does the SSS demonstrate “discriminant” validity? Correlations with the STAIC, which speak to the “convergent” validity of the SSS, unfortunately provide only half of the story.

Path analysis. To examine further the construct validity of the SSS, the authors tested a series of causal models using the statistical procedure, “path analysis.” In three separate models, the four sources of stress (and several other variables) were used to predict each of the three manifestations of stress.

I had considerable difficulty with these analyses. First, the authors do not clearly demonstrate the relevance of path analysis to the question of construct validity. Consequently, the import of these results remains unclear. Second, the logic of the general model rests on tacit--and highly questionable--assumptions regarding causality. For example, academic stress is presented as a cause of academic self-concept and academic achievement. In my view, the opposite assertion is equally plausible.

Third, no guidance is offered for interpreting these partial regression coefficients. Nor are we told that R^2 --a fundamental summary statistic in path analysis--ranges from .17 to .39 across the three analyses. These modest values, which I derived from available information, would seem to carry important implications for the tenability of the model and, more specifically, the authors' premise that sources of stress have a causal influence on manifestations of stress.

Finally, some of the authors' interpretations are questionable: (a) conclusions appear to be made about interactions among variables when, in fact, the statistical analysis did not allow for interactive effects; and (b) statements about relations among variables are based on exceedingly low effects.

CONCLUSIONS. The SSS is an easily administered and scored instrument for assessing students' perceptions of sources and manifestations of school-related stress. However, questions remain about the adequacy both of the norms and of the information pertaining to reliability and validity. Consequently, the SSS presently would appear to be more appropriate for researchers than for practitioners, insofar as the former group would be less inclined to use the SSS as a basis for forming judgments or making decisions about students.

REVIEWER'S REFERENCE

Spielberger, C. D., Edwards, C. D., Lushene, R. E., Montuori, J., & Platzek, D. (1973). *Preliminary manual for the State-Trait Anxiety Inventory for Children*. Palo Alto, CA: Consulting Psychologists Press.

Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed.; pp. 13-103). New York: Macmillan Publishing Co.