Test-Retest Reliability of a Short Form of the Children's Social Desirability Scale for Nutrition and Health-Related Research

Patricia H. Miller, PhD[1], **Suzanne D. Baxter, PhD, RD, LD, FADA, FAND**[2], **David B. Hitchcock, PhD**[3], **Julie A. Royer, MSPH**[2,4], **Albert F. Smith, PhD, MS**[5], and **Caroline H. Guinn, RD, LD**[2]

[1]Department of Psychology, San Francisco State University, San Francisco, CA

[2]Institute for Families in Society, College of Social Work, University of South Carolina, Columbia, SC

[3]Department of Statistics, University of South Carolina, Columbia, SC

[4]South Carolina Budget and Control Board, Division of Research and Statistics, Columbia, SC

[5]Department of Psychology, Cleveland State University, Cleveland, OH

Abstract

Objective—To examine test-retest reliability and internal consistency of the Children's Social Desirability Short (S-CSD) scale, consisting of 14 items from the Children's Social Desirability scale.

Methods—The previously validated S-CSD scale was classroom administered to 97 fourth-grade children (80% African American, 76% low socioeconomic status) in 2 sessions a month apart. Each classroom administration lasted approximately 5 minutes.

Results—The S-CSD scale showed acceptable levels of test-retest reliability (0.70) and internal consistency (0.82 and 0.85 for the first and second administrations, respectively). Reliability was adequate within subgroups of gender, socioeconomic status, academic achievement, and body mass index percentile. Levels of social desirability did not differ across subgroups.

Conclusions and Implications—Social desirability bias is a potential source of systematic response error in children's self-report assessments of nutrition and health-related behaviors. The S-CSD scale may be used with diverse groups of children to reliably and efficiently assess social desirability bias.

Address for correspondence: Patricia H. Miller, PhD, Department of Psychology, San Francisco State University, San Francisco, CA 94132. Phone: (415) 338-6197; Fax: (415) 338-2398; phmiller@sfsu.edu.

Keywords

social desirability; children; test-retest reliability; assessment; health

INTRODUCTION

A challenging methodological issue for nutrition and health researchers is to ascertain whether research participants provide accurate answers rather than those that they think they should provide. In particular, some participants may tend to answer questions about sensitive topics, such as eating unhealthy foods or overeating, in a socially desirable way. *Social desirability* is defined as reporting that one *never* performs a behavior that most people perform at least occasionally or *always* performs a behavior that most people usually perform but omit occasionally.[1] Questionnaires—most notably the Marlowe-Crowne Social Desirability scale—have been developed to detect social desirability response bias in adults.[2] Social desirability, assessed in this way, was correlated positively with dietary reporting errors in adults.[3] Children may also exhibit social desirability bias, which could affect their self reports about what they have eaten recently and how much they have eaten.[4,5] In fact, programs to improve child nutrition and reduce obesity that may promote children's awareness of "junk food" could encourage socially desirable answers from children concerning their food intake.

The Children's Social Desirability (CSD) scale, with slightly different versions for younger (grades 3-5) and older (grades 6-12) children, was developed in the 1960s.[6] The CSD scale for grades 3-5 has 46 yes-no items;[6] examples of items are "Do you always listen to your parents?" and "Do you ever get angry?" with a "yes" and "no" response, respectively, on these items keyed as the socially desirable response. Internal consistency, measured with Spearman-Brown-corrected split-half reliability, for the CSD scale ranged from 0.82 to 0.95 for subsamples of boys and girls in various grades, and 1-month test-retest reliability was 0.90.[6] The CSD scale, at the time it was developed, exhibited some interesting correlations with personal attributes and behaviors. Social desirability scores were higher, indicating a greater tendency to answer in a socially desirable manner, for younger children than for older children, for girls than for boys, for lower IQ children than for higher IQ children, for African American children than for European American children, and for children with lower than with higher academic achievement scores.[6,7]

In school-based research, limits on how long children can be removed from instructional class time, as well as children's limited patience or attention, may constrain the amount of data that nutrition and health researchers can collect. Thus, it may not be possible to administer the full CSD scale. Consequently, there is a need for a short scale with adequate psychometric properties. Short forms of the Marlowe-Crowne Social Desirability scale have been developed for adults.[8] Short forms based on the version of the CSD scale for older children (grades 6-12) and younger children (grades 3-5) also have been developed on samples from grades 6-9[9] and grade 4,[10] respectively. A short form for younger children is particularly important because they tend to have higher social desirability scores,[7,9] may have more trouble than older children maintaining their attention for long forms, and are the

target of many contemporary health intervention programs. External validity of the short form for younger children in nutrition research was shown in a 2010 publication in which fourth graders' social desirability scores were inversely related to their dietary reporting accuracy for energy intake at school meals (compared to direct observations of those meals).[5] This short form for younger children was developed in 2004 by drawing 14 items from the long form of the 46-item CSD scale.[10] However, as these 14 items were not administered by themselves in the 2004 study, but as part of the full 46-item scale, the short form's test-retest reliability as a stand-alone scale is unknown.

The short form in the 2004 study[10] was developed as follows: First, the full 46-item CSD scale was administered individually by telephone to 100 fourth-grade children in a school district in Georgia. Internal consistency (Cronbach's alpha) was 0.88 and 0.93 for the first and second administrations, respectively, and 1-month test-retest reliability of 0.79 (Pearson's correlation coefficient) was adequate. Then, a factor analysis of the 46-item CSD scale data showed that the proportion of common variance accounted for by the first factor was vastly greater than that of subsequent factors. Thus, the 46-item scale had a common construct from which one could select items for a short scale measuring the same construct. Next, items were selected for the short form that had high loadings on the first factor during both administrations and had non-extreme endorsement rates (required for an item to contribute to differentiation among children). Moreover, the set of items had the same proportion of items keyed "yes" for social desirability as in the full scale. The 14 items selected, when administered within the full 46-item CSD scale, had a test-retest reliability of 0.83. No gender differences were found in social desirability, and 56.9% and 56.7% of responses on the first and second administrations, respectively, were keyed as socially desirable. The current study was the next step in the development of the short form.

The main purpose of the current study was to assess test-retest reliability and internal consistency (Cronbach's alpha) of the 14 items as a stand-alone scale, hereafter referred to as the S-CSD (for Children's Social Desirability Short) scale. A new sample of fourth-grade children was tested. The present study examined the sample overall and subgroups of children to determine whether social desirability scores and test-retest reliability vary by gender, socioeconomic status (SES), child academic achievement, or BMI percentile. It is useful to know, for example, if test-retest reliability is adequate for subgroups of children who may be of particular interest to nutrition and health researchers. In addition to requiring less time to complete than the full scale, this S-CSD short form contributes several other important advances. In the current study, the short form was administered to children in a group (a class in a classroom) rather than individually, which increases the efficiency of collecting social desirability information in school-based projects. Moreover, the sample of children was more diverse than the Midwest 1960s sample on which the CSD scale was developed, which was mostly European American, of middle-to-upper-middle SES, and from a set of schools that included no large metropolitan schools.[6] In contrast, the current sample was predominantly African American, of lower-to-middle SES, and from schools in a large southern metropolitan school district. This point is important, because such populations typically are overrepresented for nutrition- and health-related problems, including obesity, and thus are the focus of much research on health disparities. Currently, there is almost no information about children's social desirability biases in these

populations. It is important to know if test-retest reliability pertains to a broad spectrum of children in society. Finally, because the many social and health changes since the 1960s could impact children's social desirability scores, it is important to examine social desirability in a contemporary sample.

METHODS

Participants

The children were from all 6 fourth-grade classrooms from 2 schools in a large metropolitan school district in South Carolina. This age was appropriate for several reasons. By this age, children can read well enough to complete a questionnaire (though research staff read the items to the children). Also, by this age children are aware of social attitudes about many health-related topics, such as which foods are considered "good" and "bad."[11] Finally, this is generally the youngest age at which children are asked to provide self-reports of their dietary intake.[12] The study had University of South Carolina Institutional Review Board approval. Parents and children provided written consent and assent, respectively.

Instruments and Measures

Table 1 lists the 14 items that comprise the S-CSD scale, along with their responses keyed as socially desirable. Each S-CSD form was scored by research staff and by computer, with 1 point for each answer keyed as socially desirable. S-CSD scale scores ranged from 0 to 14, with higher scores indicating a greater tendency to answer in a socially desirable manner.

Children's academic achievement was quantified by summing scores on the English Language Arts and Mathematics scales of the Palmetto Assessment of State Standards, the academic assessment test used by the state of South Carolina.[13] Children were classified as low SES if they were eligible for free or reduced-price school meals, and as high SES otherwise. (Children from families with income less than 130% of the U. S. Department of Agriculture, Food, and Nutrition Service poverty level were eligible for free meals and those from families with income between 130 and 185% of the poverty level were eligible for reduced-price meals.) Academic achievement scores and eligibility information are collected by the state's Department of Education and housed in the state's Division of Research and Statistics data warehouse. The Department of Education granted the Division of Research and Statistics permission to link data sets, conduct analyses, and provide aggregate results to researchers. To calculate BMI, children's weight and height were measured by research staff in the morning immediately after the first S-CSD form administration, using established procedures.[14,15] Age/gender BMI charts were used to determine BMI percentiles.[16]

S-CSD Scale Administration

In each classroom, 2 research staff distributed paper S-CSD forms, read each question aloud while the children followed along, asked children to circle "yes" or "no" on their forms, and ensured that each question was answered with only 1 answer. The S-CSD scale was administered again to the same children, in the same manner by the same research staff, 27 to 30 days later. Each classroom administration of the S-CSD scale lasted approximately 5 minutes.

Data Analysis

Pearson correlation coefficients quantified test-retest reliabilities. Paired *t*-tests were used to compare scores between administrations and z-tests compared reliabilities of subgroups of children (e.g., boys versus girls). Cronbach's alpha assessed internal consistency. A multiple linear regression model tested gender, SES, academic achievement, and BMI percentile as predictors of S-CSD scale scores. Administration-1 S-CSD scale scores were regressed on these variables and their first-order interactions. (Administration-1 scores were used because, ordinarily in practice, researchers would administer the S-CSD scale only once.) Gender and SES were dichotomous predictors, and academic achievement and BMI percentile were continuous predictors. Residual plots showed that the assumptions of normality and constant variance were justified. Statistical analyses were performed using SAS/STAT software (version 9.2, SAS Institute Inc, Cary, NC, 2002-2008).

RESULTS

The sample consisted of 97 children (45 boys, 52 girls) in the fourth grade (mean age = 10 years, 1 month). The group was 80% African American, 9% Hispanic, 7% European American, and 4% other. The mean (± standard deviation) academic achievement (Palmetto Assessment of State Standards) score was 1234.4 (± 77.0; range = 1099 to 1440), compared to a statewide mean of 1280.[17] The mean (± standard deviation) BMI percentile was 71.1 (± 27.0; range = 1.0 to 99.7).

Test-Retest Reliability and Internal Consistency

For the entire group of 97 children, 1-month test-retest reliability was 0.70 (Pearson's *r*, *P* < .0001). The scores of the 2 administrations (mean ± standard deviation: first = 6.8 ± 3.7, second = 6.7 ± 3.9) did not differ significantly (*t*-test, *P* = .68). Internal consistency (Cronbach's alpha) was 0.82 and 0.85 for the first and second administrations, respectively.

The S-CSD scale scores were also compared between subgroups of children formed by gender (boys versus girls), SES (eligible for free/reduced-price meals versus not), academic achievement (below South Carolina's state mean versus at or above state mean) and BMI percentile (below 85th percentile versus at or above 85th percentile, an expert committee's definition of "overweight or obese" for children).[18] The subgroup analyses had 92 children because the academic achievement scores of 4 children were not available and height and weight of 1 child were not obtained. Table 2 shows test-retest reliabilities (Pearson correlations) and results of z-tests (based on Fisher's z-transformation of the correlation coefficients) comparing reliabilities for children at the 2 levels of each variable. For subgroups, test-retest reliabilities ranged from 0.64 to 0.74. Test-retest reliabilities did not differ significantly by gender, SES, academic achievement, or BMI percentile (all 4 z-test adjusted *P* = 1.0). (A Bonferroni correction was used for these *P* values because the same data set was split on several variables, implying multiple tests on the same data.) Thus, the adequate test-retest reliability of the sample overall is generally reflected across subgroups.

Degree of Social Desirability

Table 2 also shows mean social desirability scores (possible range of 0 to 14) for the subgroups by administration. Scores ranged from 0 to 14, showing that the S-CSD scale differentiated among children. Overall, children chose the response keyed as socially desirable about half of the time, with 48.6% and 47.8% of such answers chosen on the first and second administrations, respectively ($P = .82$ on a paired *t*-test). Examination of the relationship of social desirability to each variable—gender, SES, academic achievement, and BMI percentile—while controlling for the other 3 variables, showed that none of the independent variables had unique significant effects (all 4 unadjusted $P > .10$). Thus, these variables did not significantly impact children's tendency to give socially desirable answers.

DISCUSSION

The S-CSD scale exhibited adequate test-retest reliability and internal consistency. Importantly, the reliabilities showed reasonable consistency for subgroups of the children formed by gender, SES, academic achievement, and BMI percentile, suggesting that the S-CSD scale is appropriate for general use. The reliabilities for the S-CSD scale were slightly less than those for the full 46-item CSD scale: test-retest reliability was 0.70 for the 14-item S-CSD scale versus 0.79 for the CSD scale given over the telephone to a similar sample, and internal-consistency was 0.82 and 0.85 for the first and second S-CSD scale administrations, respectively, versus 0.88 and 0.93 for the CSD scale.[10] This slightly reduced reliability is not surprising, because decreasing the number of items typically attenuates psychometric quality.[9]

Children's S-CSD scale scores did not differ significantly across gender, academic achievement levels, SES, and BMI percentile when controlling for the other 3 variables. In contrast, earlier studies reported higher social desirability scores by less intelligent children than by more intelligent children and by girls than by boys.[6,19] The different result regarding gender may reflect a change in gender roles since the 1960s. Consistent with this explanation, a publication in 1994[9] reported only weak evidence of gender differences. Or, differences in the sample and mode of administering the test (individually for this age by Crandall and colleagues in 1965; to a class in a classroom for the current study) may account for the different outcomes.[6] The lack of relation between S-CSD scale and BMI percentile is consistent with an earlier study with a similar sample.[5] The lack of relation between S-CSD scale and SES also is consistent with other studies.[6,19]

A limitation of this study is that the sample did not include enough children of various races to test for generality across different races or ethnic groups. Moreover, the sample included only 1 grade from 2 schools in 1 school district; thus age differences in test-retest reliability and internal consistency could not be examined.

IMPLICATIONS FOR RESEARCH AND PRACTICE

The S-CSD scale can help nutrition and health researchers and practitioners understand inaccuracies in children's reports of their nutrition and health-related behaviors. This short instrument, which can be group administered, can reliably and efficiently detect which

children may provide socially desirable responses on nutrition and health measures. Importantly, the adequate psychometric properties pertain to subgroups of particular interest to researchers studying health disparities, obesity, and nutrition.

Acknowledgments

References

1. Paulhus, DL. Measurement and control response bias. In: Robinson, JP.; Shaver, PR.; Wrightsman, LS., editors. Measures of Personality and Social Psychological Attitudes. San Diego, CA: Academic Press; 1991. p. 17-59.

2. Crowne DP, Marlowe D. A new scale of social desirability independent of psychopathology. J Consult Psychol. 1960; 24:349–354. [PubMed: 13813058]

3. Hebert JR, Peterson KE, Hurley TG, et al. The effect of social desirability trait on self-reported dietary intake measures among multi-ethnic female health center employees. Ann Epidemiol. 2001; 11:417–427. [PubMed: 11454501]

4. Baxter SD, Hardin JW, Guinn CH, Royer JA, Mackelprang AJ, Smith AF. Fourth-grade children's dietary recall accuracy is influenced by retention interval (target period and interview time). J Am Diet Assoc. 2009; 109:846–856. [PubMed: 19394471]

5. Guinn CH, Baxter SD, Royer JA, Hardin JW, Mackelprang AJ, Smith AF. Fourth-grade children's dietary recall accuracy for energy intake at school meals differs by social desirability and body mass index percentile in a study concerning retention interval. J Health Psychol. 2010; 15:505–514. [PubMed: 20460407]

6. Crandall VC, Crandall VJ, Katkovsky W. A children's social desirability questionnaire. J Consult Psychol. 1965; 29:27–36. [PubMed: 14277395]

7. Crandall VC. Personality characteristics and social and achievement behaviors associated with children's social desirability response tendencies. J Pers Soc Psychol. 1966; 4:477–486. [PubMed: 5972080]

8. Reynolds WM. Development of reliable and valid short forms of the Marlowe-Crowne Social Desirability Scale. J Clin Psychol. 1982; 38(1):119–125.

9. Carifio J. Sensitive data and students' tendencies to give socially desirable responses. J Alcohol Drug Educ. 1994; 39:74–84.

10. Baxter SD, Smith AF, Litaker MS, Baglio ML, Guinn CH, Shaffer NM. Children's social desirability and dietary reports. J Nutr Educ Behav. 2004; 36:84–89. [PubMed: 15068757]

11. Noble C, Corney M, Eves A, Kipps M, Lumbers M. Food choice and school meals: Primary schoolchildren's perceptions of the healthiness of foods and the nutritional implications of food choices. Int J Hosp Manag. 2000; 19(4):413–432.

12. Frank GC. Taking a bite out of eating behavior: Food records and food recalls of children. J School Health. 1991; 61:198–200. [PubMed: 1943041]

13. Palmetto Assessment of State Standards. [October 30, 2013] South Carolina State Department of Education. http://ed.sc.gov/agency/programs-services/45/

14. Lohman, TG.; Roche, AF.; Martorell, R. Anthropometric Standardization Reference Manual. Champaign, IL: Human Kinetics Books; 1988.

15. Maternal and Child Health Bureau. [July 22, 2013] U S Dept of Health and Human Services. Accurately weighing & measuring: Technique. http://depts.washington.edu/growth/module5/text/intro.htm

16. Kuczmarski RJ, Ogden CL, Grummer-Strawn LM, et al. CDC growth charts: United States. Advance Data. 2000; 314:1–27. http://www.ncbi.nlm.nih.gov/pubmed/11183293. [PubMed: 11183293]

17. Palmetto Assessment of State Standards. [October 30, 2013] South Carolina State Department of Education. http://ed.sc.gov/data/pass/2011

18. Barlow SE. Expert committee recommendations regarding the prevention, assessment, and treatment of child and adolescent overweight and obesity: Summary Report. Pediatrics. 2007; 120(Supplement 4):S164–S192. [PubMed: 18055651]

19. Klein EB, Gould LJ, Corey M. Social desirability in children: An extension and replication. J Consult Clin Psychol. 1969; 33:128. [PubMed: 4887401]

PRACTICE POINTS

The long form of the Children's Social Desirability (CSD) scale has 46 items; the Children's Social Desirability (S-CSD) Short scale is a 14-item subset of the 46 items.

The 14-item S-CSD scale, which can be group administered, is appropriate for general use and with diverse groups (by gender, SES, academic achievement, BMI percentile).

The 14-item S-CSD scale is a reliable and efficient way to detect which children may provide socially-desirable responses on nutrition and health-behavior assessments.

Table 1

Items on the Children's Social Desirability Short (S-CSD) Scale[a]

Item	Socially Desirable Response
1. Have you ever felt like saying unkind things to a person?	No
2. Are you always careful about keeping your clothing neat and your room picked up?	Yes
3. Do you sometimes feel like staying home from school even if you are not sick?	No
4. Do you ever say anything that makes somebody else feel bad?	No
5. Are you always polite, even to people who are not very nice?	Yes
6. Sometimes, do you do things you've been told not to do?	No
7. Do you always listen to your parents?	Yes
8. Do you sometimes wish you could just play around instead of having to go to school?	No
9. Have you ever broken a rule?	No
10. Do you sometimes feel angry when you don't get your way?	No
11. Do you sometimes feel like making fun of other people?	No
12. Do you always do the right things?	Yes
13. Are there some times when you don't like to do what your parents tell you?	No
14. Do you sometimes get mad when people don't do what you want them to do?	No

[a]These 14 items from the S-CSD scale are from Baxter et al.,[10] who selected them from the 46-item CSD scale of Crandall et al.[6]

Table 2

Children's Social Desirability Short (S-CSD) Scale Means (and Standard Deviations [SD]) and Test-Retest Reliabilities of Subgroups

Variable	n	Mean (SD)[a] 1st Administration	Mean (SD) 2nd Administration	Pearson Correlation	z-test P value[b]
Gender					
Girls	50	6.5 (3.8)	6.7 (3.9)	.66	1
Boys	42	7.2 (3.6)	6.8 (3.9)	.71	
Socioeconomic status					
Free/reduced price meals	70	7.2 (3.5)	7.3 (3.7)	.69	1
Full price meals	22	5.5 (3.8)	4.9 (3.9)	.60	
Academic achievement					
< State mean (1280)	65	7.1 (3.6)	7.0 (3.7)	.65	1
State mean	27	6.1 (3.8)	6.1 (4.2)	.74	
Body mass index percentile					
< 85th percentile	50	6.7 (3.8)	6.4 (3.9)	.72	1
85th percentile	42	7.0 (3.6)	7.1 (3.9)	.63	
Total	97[c]	6.8 (3.7)	6.7 (3.9)	.70	

[a] Possible scoring range for S-CSD scale is 0-14.

[b] P values for the 4 z-tests comparing correlations between subgroups were adjusted using the Bonferroni method (4 unadjusted P values = .66, .55, .47, and .36, respectively).

[c] The total n was 97 but the subgroup n was 92 because academic achievement scores were not available on 4 children and height and weight were not obtained on 1 child.