Published in final edited form as:

Pers Individ Dif. 2015 September ; 83: 85–90. doi:10.1016/j.paid.2015.03.039.

Children's Social Desirability: Effects of Test Assessment Mode

Patricia H. Miller^a, Suzanne D. Baxter^b, Julie A. Royer^{b,c}, David B. Hitchcock^d, Albert F. Smith^e, Kathleen L. Collins^b, Caroline H. Guinn^b, Alyssa L. Smith^b, Megan P. Puryear^b, Kate K. Vaadi^b, and Christopher J. Finney^{b,c}

Patricia H. Miller: phmiller@sfsu.edu; Suzanne D. Baxter: sbaxter@mailbox.sc.edu; Julie A. Royer: royerj@mailbox.sc.edu; David B. Hitchcock: hitchcock@stat.sc.edu; Albert F. Smith: a.f.smith@csuohio.edu; Kathleen L. Collins: colli239@mailbox.sc.edu; Caroline H. Guinn: cguinn@mailbox.sc.edu; Alyssa L. Smith: smith335@mailbox.sc.edu; Megan P. Puryear: pricev@mailbox.sc.edu; Kate K. Vaadi: vaadi@mailbox.sc.edu; Christopher J. Finney: finneyc@mailbox.sc.edu

^aDepartment of Psychology, San Francisco State University, San Francisco, CA, USA, 94132

^bInstitute for Families in Society, College of Social Work, University of South Carolina, Columbia, SC, USA, 29208 (Baxter, 803-777-1824 extension 12)

^cSouth Carolina Revenue and Fiscal Affairs Office, Health and Demographics Section, Columbia, SC, USA, 29201

^dDepartment of Statistics, University of South Carolina, Columbia, SC, USA, 29208

^eDepartment of Psychology, Cleveland State University, Cleveland, OH, USA, 44115

Abstract

This study examined a recently developed short version of the Children's Social Desirability (CSD-S) scale with 157 fourth-grade children. Of interest was a) whether one-month test-retest reliability would vary as a function of test assessment mode (interview or classroom), gender, race, SES, and BMI percentile, and b) whether the degree of social desirability would vary as a function of these same variables. The CSD-S scale showed good test-retest reliability for both interview and classroom assessment modes (.85 and .83, respectively). Internal consistency also was good (first interview administration = .84; first classroom administration = .81). Reliability was good and did not vary significantly over assessment mode or any child subgroup variables, suggesting that the CSD-S scale is appropriate for general use. The interview mode elicited significantly more socially desirable answers than did the classroom mode. Social desirability did not differ across child subgroups. Some of these findings were examined, and replicated, on another sample. Thus, the CSD-S scale may be used with diverse groups of children to a) reliably assess a social desirability bias that may systematically bias other self-reports of interest to researchers and b) examine individual differences in degree of social desirability.

Keywords

social desirability; children; test-retest reliability; interview or classroom assessment mode

1. Introduction

Social desirability refers to some people's tendency to present themselves in a positive light by over-reporting culturally approved positive behaviors and under-reporting negative behaviors (e.g., Crandall, Crandall, & Katkovsky, 1965; Crowne & Marlowe, 1960). Specifically, they may report that they *never* perform a behavior that most people perform at least occasionally or *always* perform a behavior that most people usually perform but omit occasionally (Paulhus, 1991). Social desirability in both adults and children is of interest for methodological and theoretical reasons, as well as for applications. Regarding methodology, a challenging issue is to ascertain whether research participants provide accurate answers rather than those approved by others. This is particularly an issue when assessing sensitive topics, such as eating unhealthy foods, overeating, or engaging in risky behaviors. It is plausible that people who over-report positive behaviors on a social desirability self-report assessment will also do so on the self-report of interest in the study. Thus, a social desirability bias makes it difficult to distinguish people with favorable traits or behaviors from those over-reporting them. Researchers can use a social desirability assessment to control for this bias when examining the self-report of interest or can eliminate participants who demonstrate a strong social desirability bias. Regarding theory, information about the tendency to report socially desirable behaviors is relevant to an understanding of individual differences and personality. Finally, practitioners can use information about social desirability to aid their interpretation of clients' behaviors.

The present study had two aims. Aim one was a methodological one—to examine one-month test-retest reliability of a recently developed short version of a social desirability scale for children, as a function of interview versus classroom assessment mode, gender, race, SES (socioeconomic status), and BMI (body mass index) percentile. Aim two was a theoretical one—to examine the degree of social desirability bias shown as a function of these variables.

Researchers have developed questionnaires, such as the widely used Marlowe-Crowne Social Desirability scale, to detect social desirability response bias in adults (Crowne & Marlowe, 1960). The Children's Social Desirability (CSD) scale was developed in the 1960s (Crandall et al., 1965), with slightly different versions for younger (grades 3–5) and older (grades 6–12) children. The CSD scale for grades 3–5 has 46 yes/no items—for example, “Do you always listen to your parents?” and “Do you ever get angry?”—with a “yes” and “no” response, respectively, keyed as the socially desirable response. More socially desirable responses were given by younger children than by older children, by girls than by boys, by lower IQ children than by higher IQ children, by African American children than by European American children, and by children with lower than with higher academic achievement scores (Crandall, 1966; Crandall et al., 1965).

There has been a need for a shorter version of this 46-item scale for research conducted at schools. Limits on the amount of time that children can be removed from class, children's limited patience and attention, and researchers' primary interest in collecting other data make administration of the full CSD scale impractical.

Short versions of the CSD scale for older children (grades 6–12) and younger children (grades 3–5) have been developed for children in grades 6–9 (Carifio, 1994) and grade 4 (Baxter et al., 2004). The short version for younger children is particularly important because they tend to have higher social desirability scores (Carifio, 1994; Crandall, 1966), may have more trouble than older children sustaining attention, and are the target of many health intervention programs. The short version for younger children (CSD-S scale) was developed by drawing 14 items from the longer 46-item CSD scale (Baxter et al., 2004). The CSD-S scale has adequate test-retest reliability (.70) and adequate internal consistency (.82) in classroom assessment for children in grade 4 (Miller et al., 2014). Good external validity was shown, as fourth graders' social desirability scores were inversely related to their dietary reporting accuracy for school meals (comparing self-reports to direct observations of those meals) (Guinn et. al., 2010).

One limitation of the CSD-S scale is that its psychometric properties have been examined only for classroom assessment. However, researchers often assess children in one-on-one interviews. Thus, the present study examined whether an interview assessment mode also provided adequate test-retest reliability for the CSD-S scale, and whether this reliability differed from that in a classroom assessment mode. Also of interest was whether the CSD-S scale showed adequate test-retest reliability for diverse subgroups of children formed by gender, race, SES, and BMI percentile, and whether their reliabilities differed. It is important to know reliabilities for subgroups of children other than middle class European Americans, because other groups often are the subject of research on health disparities. We examined BMI percentile, in addition to usual demographic variables, because of the considerable current interest in studying childhood obesity.

In addition to assessing one-month test-retest reliabilities and internal consistency in two assessment modes, the present study examined differences in the degree of social desirability as a function of assessment mode, gender, race, SES, and BMI percentile. Significant group differences would indicate that it is particularly important for researchers to include a social desirability assessment for certain populations. Moreover, any difference would increase knowledge of the contributors to individual differences in social desirability among children. Currently, there is little information about children's social desirability biases in non-white, low SES, and high BMI populations (but see Miller et al., 2014). The present study provides information about these possible group differences.

2. Method

2.1 Participants

During Spring, 2014, data were collected from 157 children (72 girls, 85 boys; mean age = 10 years, 1 month) from 21 fourth-grade classes at five schools (three urban and two rural) in four school districts in a southeastern U.S. state. This age was selected because the CSD-

S scale was developed on a sample of this age. By fourth grade, children can read well enough to complete a questionnaire in a classroom or interview setting (though research staff read items to children in both assessment modes). The group was African American (n = 79), European American (n = 65), Hispanic (n = 7), and mixed (n = 6) (as reported by parents for school records). For statistical analyses, a combined Hispanic/mixed group (n = 13) was formed due to small sample sizes. The mean ($\pm$ standard deviation) BMI percentile was 68.84 ± 29.96 . Two BMI groups were formed: BMI at or above the 85th percentile (n = 63) is an expert committee's definition of "overweight or obese" for children (Barlow, 2007); the other group (n = 94) was below the 85th percentile. The low SES group (n = 118) was eligible for free/reduced-price meals (i.e., 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). Children paying full meal price (n = 25) formed the other SES group. The study had university Institutional Review Board approval. Parents and children provided written consent and assent, respectively.

2.2 Assessments and procedure

The CSD-S scale (see Table 1) consists of 14 items from the 46-item CSD (see Baxter et al., 2004, and Miller et al., 2014, for development of the CSD-S scale and its psychometric properties). CSD-S scale scores can range from 0 to 14, with higher scores indicating a greater tendency to answer in a socially desirable manner. To calculate BMI, children's weight and height were measured twice by research staff in the morning immediately after the first CSD-S scale administration, using established procedures (Lohman, Roche, & Martorell, 1988; Maternal and Child Health Bureau of USDHHS, 2013). The Centers for Disease Control and Prevention's age/gender BMI charts were used to determine BMI percentiles (Kuczmarski et al., 2000).

2.3 CSD-S scale assessment modes and administrations

Classes within each school were randomly assigned to the interview or classroom assessment mode. For the interview assessment mode (10 classes; 77 children), one of three researchers total read each item aloud to individual children in a private location at school. Children were asked to respond "yes" or "no" verbally, and the researcher marked the answer given on the child's form. The CSD-S scale was administered again to the same children, in the same manner, by a different researcher, 28 to 32 days later. Each interview was audio-recorded. A non-interviewing researcher compared the audio-recording against the form for each child for each administration to ensure that each answer provided by the child was what the researcher had circled on the form.

For the classroom assessment mode (11 classes; 80 children), in each classroom, two of a total of three researchers distributed paper CSD-S scale forms. Researcher 1 read each item aloud while children followed along, and asked children to circle "yes" or "no" on their forms, while Researcher 2 walked around the classroom to ensure that each child answered each item with only one answer. The CSD-S scale was administered again to the same children, in the same manner, 28 to 32 days later. One researcher in the second

administration was not present in the first administration, and the Researcher 1 role was performed by a different person than in the first administration.

Weight and height were measured on the same day as, and immediately after, the first CSD-S scale administration. Because this study was part of a larger parent study, a 5-item food insecurity assessment was always given just before each administration of the CSD-S scale for both the classroom and interview assessment modes. Thus, the order was food insecurity, then CSD-S scale, and finally weight and height (with weight and height collected at the first administration only).

2.4 Data analysis

Test-retest reliability was analyzed using Pearson correlation coefficient r between test results on the first and second administrations; reliability was judged using the recommendation of Nunnally and Bernstein (1994): Reliabilities of .80 or more are fine when making decisions about groups. Also, we compared test-retest reliability across several subgroup splits—gender, SES, race, and BMI percentile (i.e., below 85th percentile or not)—using z-tests to compare two correlations. To account for the multiple dependent inferences produced by splitting the sample in several ways, we used a Holm-Bonferroni correction and controlled the family-wise Type I error rate at .10. To determine whether the degree of social desirability varied by assessment mode, race, gender, SES group, and BMI group, we fit a generalized estimated equation (GEE) model, allowing for possible correlation between responses in children from the same teacher. Effects of independent variables were tested using Wald z-tests, with the Benjamini-Hochberg method to control the false discovery rate at .10 (with this level chosen to maintain power to find significant factors even with the multiple-testing correction). Two-factor interaction effects were originally included, with the option to remove interaction terms if none were close to significant.

3. Results and discussion

Over participants, scores on the CSD-S scale ranged from 0 to 14, showing that this scale differentiates among children. Table 1 shows the percentage of children giving a socially desirable response on each item (on the first administration, because researchers typically would administer the CSD-S scale only once). Table 2 shows mean social desirability scores, and test-retest reliabilities for each assessment mode and for the demographic subgroups.

3.1 Test-retest reliabilities

The first aim concerned adequacy of one month test-retest reliability and internal consistency for each assessment mode—interview and classroom—of the CSD-S scale. For the interview mode, reliability (Pearson's r) was .85, $p < .0001$, and internal consistency (on the first administration, Cronbach's alpha) was .84, $p < .0001$. For the classroom mode, reliability was .83, $p < .0001$, and first administration internal consistency was .81, $p < .0001$. Thus, for each assessment mode, reliability and internal consistency were good.

The first aim also concerned differences in test-retest reliability and internal consistency of the CSD-S scale among various subgroups of the diverse sample. In addition to assessment mode, child subgroups were formed by gender (2), SES (low versus high), race (3), and BMI percentile (2), as described in the Method section (2.1). All child subgroups demonstrated good test-retest reliabilities (see Table 2), ranging from .81 to .89, all p -values $< .0001$.

We examined whether reliabilities differed according to assessment mode, gender, SES, race, and BMI percentile, using z -tests to compare corresponding Pearson correlations of the two tests across subgroups. No significant differences in test-retest reliabilities were found; (unadjusted) p -values were .65, .42, .50, .45, and .63, respectively. Under a Holm-Bonferroni correction for multiple testing, no comparison was significant at the .10 family-wise level. Thus, test-retest reliability appeared to be consistent across assessment mode and child subgroups, suggesting that the CSD-S scale is appropriate for general use. The adequate reliabilities for all child subgroups are consistent with a previous study (Miller et al., 2014), though the present study added race as a subgroup category.

3.2 Differences in degree of social desirability

The second aim concerned differences in the degree of social desirability (in the first administration) by assessment mode, gender, SES, race, and BMI. We fit a GEE model that allowed for possible correlation between responses in children from the same teacher. We initially fit a model including second-order interaction terms. Because none of the interaction effects were close to significant, we removed them from the model, pooling their sums of squares (SS) into the error SS (Kutner et al., 2004) and refit the model without interactions. A residual analysis showed no problems with model assumptions or fit. We tested the effects of assessment mode, gender, SES, race, and BMI on mean CSD-S scale score with Wald z -tests. A Benjamini-Hochberg correction was used to maintain the false discovery rate at .10. Given this, assessment mode had a significant effect on CSD-S scale scores (unadjusted $p = .015$; adjusted $p = .07$), with a greater mean score for the interview mode ($M = 8.43$) than for the classroom mode ($M = 6.83$). None of the other explanatory variables was significant (all adjusted p -values $> .13$) in the presence of other variables in the model. The lack of difference in social desirability across child subgroups is consistent with a previous study (Miller et al., 2014), though the present study added race as a subgroup category.

The greater social desirability in the interview mode was robust. Examination of each item (see Table 1) revealed greater scores in the interview mode on all 14 items, with the difference in the percentage of answers that indicated social desirability in the two modes ranging from 2% (36% vs. 34% on item # 6) to 29% (55% vs. 26% on item # 1). Greater social desirability in interviews may indicate that the adult interviewer's presence, and possible disapproval of a socially undesirable answer, is more salient to the child in that mode. In the one-on-one mode, the interviewer is very near, with attention focused on the child. Children may perceive the expectations of the interviewer or tester, an authority figure, differently in the two settings; a child may feel less pressure to meet adult expectations in the classroom mode. Moreover, greater social desirability scores in the

interview mode may indicate less accuracy than in the classroom mode, because it is unlikely that children always or never perform the behaviors mentioned.

The lack of effect of gender, SES, or BMI is consistent with recent studies with a similar sample (Guinn et al., 2010; Miller et al., 2014). In contrast, earlier studies reported greater social desirability scores by girls than by boys (Crandall et al., 1965; Klein, Gould, & Corey, 1969). The different results regarding gender may reflect a change in gender roles since the 1960s or differences in the samples. Societal change or sample differences may also account for our finding of no effect of race in comparison to earlier studies (Crandall, 1966; Crandall et al., 1965). The lack of relation between the CSD-S scale and SES is also consistent with other studies (Crandall et al., 1965; Klein et al., 1969; Miller et al., 2014).

3.3 Replication

We were able to examine the replicability and generalizability of our results, as well as obtain additional information about the stability of individual differences in social desirability, in an unpublished data set with a different sample and slightly different design. The 56 fourth-grade children (28 girls, 28 boys) also were racially diverse (36 African American, 9 European American, 6 Hispanic, 5 mixed), and from 11 classes at the same schools as in the main study. The procedures for administering the interview and classroom CSD-S scale assessment modes were virtually identical to the main study. This sample a) provided a replication of the test-retest reliability with the classroom assessment mode, and b) permitted us to examine correlations between social desirability scores on the interview and classroom assessment modes, because all children received both modes. This contrasts with the design of our main study with the comparison between children (i.e., children assigned to either the interview or classroom modes but not both). These 56 children had three CSD-S scale administrations—interview assessment (during a one-on-one interview consisting of [listed in order given] dietary recall, food insecurity survey, CSD-S scale, and weight and height measurements), followed two to three-and-a-half months later by the first classroom CSD-S scale administration, followed one month later by a second classroom CSD-S scale administration.

Replicating the good test-retest reliability of classroom assessments in the main study, the two classroom assessments had a Pearson correlation of .82, $p < .0001$. Cronbach's alpha was identical (.80) for the interview assessment and the first classroom administration, showing good alphas for both versions. Moreover, the Pearson correlation between the interview and first classroom administration was .64, $p < .0001$, which is new information made possible by repeated measures on assessment mode. Thus, even though social desirability scores were greater on the interview than classroom assessment mode in both the main study (comparing between-children scores) and this replication study (comparing within-children scores)—interview $M = 8.50$, classroom $M = 7.64$, $p = .002$ —children were somewhat consistent in their ordering within the group over three months. One implication is that researchers interested in individual differences in children's social desirability could use either assessment mode.

The repeated-measures design also extended information about item differences reported in Table 1 for the main study. Supporting results of the main study, more children gave the

socially desirable response in the interview than in the classroom assessment mode for 12 of 14 items when comparing Session #1 Interview to Session #2 Classroom, and also for 12 of 14 items when comparing Session #1 Interview to Session #3 Classroom (see Table 3).

In summary, this study replicated the good test-retest reliability and assessment mode differences in the main study despite differences in the context of assessment. Specifically, the replication study differed from the main study in that the interview assessment also included assessments of dietary recall (the focus of that study). Also, when children completed the CSD-S scale classroom assessments, they had already completed the CSD-S scale interview assessment. Thus, researchers can still expect good test-retest reliability when the CSD-S scale is given in the context of other assessments.

3.4 General discussion

Regarding methodology, these results, together, suggest that researchers could rely on either assessment mode of the CSD-S scale, because both have good test-retest reliability and internal consistency. The earlier report of adequate test-retest reliability of the CSD-S scale in a classroom setting (Miller et al., 2014) is extended to a new sample and, in fact, the classroom testing in the current study shows higher reliability (.83) than the earlier sample (.70). The good reliability of the interview assessment mode is new information. Importantly, reliabilities in both assessment modes showed reasonable consistency for diverse subgroups of children formed by gender, SES, race, and BMI, indicating wide applicability.

Regarding implications for our understanding of individual differences in social desirability, the results suggest that children's degree of social desirability is consistent over time, at least for a one-month interval. However, context—one-on-one interview versus a group setting—influences the degree of social desirability that fourth-grade children express. The greater number of socially desirable answers in the interview assessment mode than the classroom mode suggests that the consistent social desirability bias on which children vary is affected by their surroundings. Still, the highly significant correlation between the social desirability scores with the two testing modes on the same children (repeated measures) in the replication study suggests that the individual differences are somewhat similar across the two assessment modes. Moreover, the similarity, rather than difference, in consistency and level of social desirability across children varying in gender, SES, race, and BMI, suggests that the individual differences are not due to these factors. The strength of a social desirability bias may instead reflect both a psychological characteristic of the child and the assessment context.

The main study and its replication have several limitations. The samples included only fourth-grade children in one U.S. state. Also, the sample did not include enough children of various races to test for generality across several races or ethnic groups. As social desirability was not the focus of the larger parent study, future research could more solidly examine social desirability assessment mode using four groups, two with interview and two with classroom mode; in the retest, one of the two groups would use the same mode, and the other group would use the other mode.

4. Conclusions

The present study is the first to provide test-retest reliabilities and mean performance on a short social desirability instrument for both one-on-one and classroom assessment modes, across diverse subgroups of children. The CSD-S scale fills researchers' need for a short, reliable test that can be used to detect a social desirability bias in diverse samples of children. The test can be used flexibly—in group or individual interview assessments, and in the context of other instruments. Because the classroom assessment is more efficient in terms of time and money, it is important to know that test-retest reliability is good in that setting. In addition to this methodological contribution, the study provides new information about social desirability as an individual difference. Individual differences in social desirability appeared in all subsamples and in both assessment modes. Interestingly, its prevalence does not differ according to race, gender, SES, or BMI, which is of particular interest to researchers studying health disparities, obesity, and nutrition. The fact that more socially desirable responses are elicited in a one-on-one interview than in the classroom informs theory by showing that this tendency is affected by the context and informs methodology by showing that assessment mode must be considered when degree of social desirability is of interest.

Acknowledgments

This research was supported by grant R01 HL103737 from the National Heart, Lung, and Blood Institute of the National Institutes of Health. The second author was Principal Investigator. The content of this article is solely the responsibility of the authors and does not necessarily represent the official views of the National Heart, Lung, and Blood Institute of the National Institutes of Health, or the South Carolina Revenue and Fiscal Affairs Office.

The authors thank the children and staff of the five elementary schools in four school districts — Richland One, Lexington Three, Lexington Four, and Lexington-Richland Five — in South Carolina, U.S., for allowing data collection.

References

- 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–S192.10.1542/peds.2007-2329C [PubMed: 18055651]
- Baxter SD, Smith AF, Litaker MS, Baglio ML, Guinn CH, Shaffer NM. Children's social desirability and dietary reports. *Journal of Nutrition Education and Behavior*. 2004; 36:84–89.10.1016/S1499-4046(06)60138-3 [PubMed: 15068757]
- Carifio J. Sensitive data and students' tendencies to give socially desirable responses. *Journal of Alcohol and Drug Education*. 1994; 39:74–84.
- Crandall VC. Personality characteristics and social and achievement behaviors associated with children's social desirability response tendencies. *Journal of Personality and Social Psychology*. 1966; 4:477–486.10.1037/h0023891 [PubMed: 5972080]
- Crandall VC, Crandall VJ, Katkovsky W. A children's social desirability questionnaire. *Journal of Consulting Psychology*. 1965; 29:27–36.10.1037/h0020966 [PubMed: 14277395]
- Crowne DP, Marlowe D. A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology*. 1960; 24:349–354.10.1037/h0047358 [PubMed: 13813058]
- 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. *Journal of Health Psychology*. 2010; 15:505–514.10.1177/1359105309353814 [PubMed: 20460407]

- Klein EB, Gould LJ, Corey M. Social desirability in children: An extension and replication. *Journal of Consulting and Clinical Psychology*. 1969; 33:128.10.1037/h0027339 [PubMed: 4887401]
- Kutner, M.; Nachtsheim, C.; Neter, J.; Li, W. *Applied linear statistical models*. 5. Boston: McGraw-Hill; 2004.
- Lohman, TG.; Roche, AF. *Anthropometric standardization reference manual*. Martorell, R., editor. Champaign, IL: Human Kinetics Books; 1988.
- Miller PH, Baxter SD, Hitchcock DB, Royer JA, Smith AF, Guinn CH. Test-retest reliability of a short form of the children's social desirability scale for nutrition and health-related research. *Journal of Nutrition Education and Behavior*. 2014; 46(5):423–428.10.1016/j.jneb.2013.11.002 [PubMed: 24418615]
- Paulhus, DL. Measurement and control of 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.

Web references

- Kuczmarski, RJ.; Ogden, CL.; Grummer-Strawn, LM.; Flegal, KM.; Guo, SS.; Wei, R.; Johnson, CL. *Advance Data from Vital and Health Statistics*. Vol. 314. Hyattsville, MD: National Center for Health Statistics; 2000. CDC growth charts: United States; p. 1-27.<http://www.ncbi.nlm.nih.gov/pubmed/11183293> [Accessed March 4, 2014]
- Maternal and Child Health Bureau. U. S. Dept. of Health and Human Services. [Accessed March 4, 2014] Accurately weighing & measuring: Technique. <http://depts.washington.edu/growth/module5/text/intro.htm>

Highlights

- We examined a short version of the Children's Social Desirability scale (CSD-S).
- Test-retest reliability and internal consistency of the CSD-S were good.
- Reliability and consistency were good in interview and classroom assessment modes.
- Good reliability generalized across child subgroups of race, gender, SES, and BMI.
- Interview mode elicited more socially desirable responses than classroom mode.

Table 1

Percentage of children giving socially desirable response per item on the Children's Social Desirability Short (CSD-S) scale in the main study

CSD-S scale items	Socially desirable response	Percentage of children giving socially desirable response	
		One-on-one interview assessment	Classroom assessment
1. Have you ever felt like saying unkind things to a person?	No	54.6	26.3
2. Are you always careful about keeping your clothing neat and your room picked up?	Yes	89.6	85.0
3. Do you sometimes feel like staying home from school even if you are not sick?	No	41.6	36.3
4. Do you ever say anything that makes somebody else feel bad?	No	74.0	71.3
5. Are you always polite, even to people who are not very nice?	Yes	81.8	63.8
6. Sometimes, do you do things you've been told not to do?	No	36.4	33.8
7. Do you always listen to your parents?	Yes	79.2	67.5
8. Do you sometimes wish you could just play around instead of having to go to school?	No	45.5	32.5
9. Have you ever broken a rule?	No	46.8	30.0
10. Do you sometimes feel angry when you don't get your way?	No	44.2	30.0
11. Do you sometimes feel like making fun of other people?	No	81.8	75.0
12. Do you always do the right things?	Yes	61.0	45.0
13. Are there some times when you don't like to do what your parents tell you?	No	41.6	35.0
14. Do you sometimes get mad when people don't do what you want them to do?	No	64.9	51.3

Table 2

Children's Social Desirability Short Scale (CSD-S) means (and standard deviations) and test-retest reliabilities of subgroups

Variable	n	Mean (SD) ^a 1 st administration	Mean (SD) 2 nd administration	Pearson correlation
Assessment mode				
Interview	77	8.4 (3.7)	8.2 (4.1)	.85
Classroom	80	6.8 (3.5)	7.1 (3.9)	.83
Gender				
Girls	72	8.2 (3.6)	8.2 (3.8)	.82
Boys	85	7.1 (3.7)	7.2 (4.2)	.86
Socioeconomic status ^b				
Free/reduced price meals	118	8.0 (3.6)	8.2 (4.0)	.82
Full price meals	25	6.7 (3.9)	6.6 (4.0)	.87
Race				
European American	65	7.3 (3.7)	7.1 (4.2)	.86
African American	79	7.8 (3.6)	7.9 (3.7)	.81
Hispanic/mixed	13	8.4 (4.6)	8.5 (5.0)	.89
Body mass index percentile				
< 85 th percentile	94	7.2 (3.7)	7.3 (3.7)	.83
85 th percentile	63	8.2 (3.7)	8.1 (4.1)	.86
Total	157	7.6 (3.7)	7.6 (4.0)	.84

^a Possible scores for CSD-S scale range from 0 to 14.

^b The total n was 157 but the SES subgroups n was 143 because it was not possible to link 14 children to the aggregate results of the data provided by the state's Department of Education.

Table 3

Percentage of children giving socially desirable response per item on the Children's Social Desirability Short (CSD-S) scale in the replication study

CSD-S scale item ^a	Percentage of children giving socially desirable response		
	Session #1 One-on-one interview assessment	Session #2 1 st administration Classroom assessment	Session #3 2 nd administration Classroom assessment
1	48.2	46.4	35.7
2	91.1	83.9	83.9
3	42.9	33.9	35.7
4	69.6	71.4	58.9
5	75.0	66.1	67.9
6	41.1	44.7	39.3
7	83.9	75.0	66.1
8	48.2	32.1	35.7
9	50.0	44.6	48.2
10	41.1	37.5	42.9
11	85.7	75.0	69.6
12	58.9	55.4	51.8
13	51.8	46.4	46.4
14	62.5	51.8	62.5

^aSee Table 1 for CSD-S scale items and corresponding socially desirable response for each.