DOCUMENT RESUME

ED 343 915 TM 018 046

AUTHOR Carifio, James
TITLE Parallel Short Forms of the Crandall Social Desirability Test for Children: Shortening Instruments for Research Purposes.
PUB DATE Mar 92
NOTE 19p.; Paper presented at the Annual Meeting of the Eastern Educational Research Association (15th, Hilton Head, SC, March 5-9, 1992).
PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Attitude Measures; Elementary School Students; Factor Analysis; Intermediate Grades; Junior High Schools; Junior High School Students; Middle Schools; Program Evaluation; *Rating Scales; *Research Methodology; Sampling; Social Desirability; *Test Construction; Testing Problems; Test Length; Test Validity
IDENTIFIERS *Crandall Social Desirability Scale; *Parallel Test Forms; Test Retest Reliability; Test Shortening Procedures

ABSTRACT

Researchers and program evaluators would often like to use a particular instrument, but do not because it is too long or would require too much testing time. Having a validated set of objective procedures for reducing the size of an instrument could improve many research and evaluation efforts. This paper reports the results of test reduction or shortening procedures devised and used in the development of parallel short forms of the Crandall Social Desirability Test for Children (CSDTC), a 48-item scale that measures a person's tendency to give socially desirable answers to questions rather than the subject's true views or feelings. The first procedure was to draw four random samples of 12 items (25% of the CSDTC) to produce parallel short forms based on a rationale from sampling theory. These samples were assessed through responses of 356 fifth and sixth graders in a Boston (Massachusetts) middle school to the full scale and shortened versions. A factor analytic approach was also used to make parallel short forms by taking the four items that loaded most highly on each of the three major factors from the full scale administration. Cross validation with 181 sixth and eighth graders; a test-retest reliability study with 55 seventh graders; and a validity study with 254 seventh, eighth, and ninth graders supported the short form derived through factor analysis. Results demonstrate empirically that long instruments can be effectively reduced, and that factor analytic strategies are superior to sampling strategies for this purpose. Six tables present form items and results from the validity studies. Four references are listed. (SLD)

Parallel Short Forms of the Crandall Social Desirability Test for Children: Shortening Instruments for Research Purposes

James Carifio
University of Massachusetts at Lowell

Abstract

The present study empirically shows that long instruments can be effectively reduced for research purposes in a manner that will preserve their psychometric properties and construct validity by either random sampling or factor analytical techniques. The present study, however, also empirically shows that factor analytical instrument reduction strategies and techniques are superior to random sampling strategies and techniques and should be used when the appropriate data is available. These findings, therefore, should help researcher and program evaluators to improve the complexity, sophistication and quality of their designs and studies by employing reduced versions of measurement instrument when and where appropriate.

Two parallel short forms of the Crandall Social Desirability Scale for Children were developed and validated in the present study using factor analytical instrument reduction techniques. These two parallel short forms should be of great use to many researchers, but particularly to those who do research on sensitive and/or controversial topics such as alcohol and drug usage by questionnaire or interview techniques. Both implications and theory are discussed and several practical recommendations made.

Paper presented at the annual conference of the Eastern Educational Research Association, Hilton Head, So. Carolina, March 5-9, 1992.

Today, researchers and program evaluators tend to gather data on several variables as part of their overall study design rather than data on just a single dependent variable. However, in this context of increased complexity, researchers and program evaluators often would like to use a particular test or scale in their experiment or evaluation, but refrain from doing so because the test or scale is too long and would require too much testing time and/or might possibly fatigue subjects when the time to collect the data on all of the variables included in the study is taken into consideration. A strong or stronger design, therefore, often is not possible because of the total testing time and test length problems.

The above practical research problem very often impedes researchers and program evaluators because the research methodology literature is relatively (if not completely) silent on strategies that may be employed to shorten standardized tests and scales for research and evaluation purposes, while maintaining the test's or scale's reliability and validity and thus interpretability. Having a validated set of objective instrument reduction procedures, therefore, would unequivocably improve the quality of many research and program evaluation efforts and activities.

Purpose

The purpose of this paper is to report the results of the test reduction or shortening procedures that were devised and used in the development of parallel short forms of the Crandell Social Desirability Test for Children (CSDTC). The CSDTC is a 48 item scale that measures a person's tendency to give social desirable (or acceptable) answers to questions rather than answers that express the person's true views, opinions, or feelings. This tendency often introduces a response bias into the data that makes clear and unambiguous interpretation of the data and it's meaning difficult if this tendency is not controlled through measurement and data adjustments in the data analysis process.

One context where the tendency to give social desirable responses strongly affects data quality, interpretatation and analysis is the situation where data is collected on sensitive and/or controversial topics such as alcohol and drug usage by questionnaire or interview techniques. This context almost always introduces an element of response bias into the data, even when the data are collected anonymously or by anonymity preserving techniques, particularly with adolescents and children. Other situations where the tendency to give social desirable responses affects data quality is when subjects give answer to questions about their social views and beliefs. In this context the influences of a tendency to give social desirable responses can be quite subtle. Having parallel short forms of the CSDTC to embedded

in questionnaire and interview protocols on social views, beliefs and sensitive and/or controversial topics, therefore, would be of great benefit to a broad range or researchers and program evaluators, particularly as questionnaires and iterviews on these types on variables tend to be long.

Methodology

Two very different instrument reduction strategies and techniques were employed to produce parallel short forms of the CSDTC. The first procedure used was to draw four 25% random samples of items from the 48 item long CSDTC form to produce parallel short forms. The rationale for this procedure is straight from sampling theory. The standard error of sampling from a finite population (of items) may be used to estimate a person's long form score, and the general range of the correlations between short forms and short forms and long forms also may be estimated. These particular points are inherent in Cronbach's (1951) conceptual view and theory of his coefficient alpha, which estimates the upper limits these correlations. Three further points also need to be noted here.

First, when one is doing empirical research (i.e., group as opposed to individual measurement and assessment) one may employ estimates of variables and estimates that are somewhat less reliable as one is making decisions about hypotheses and relationsips and not specific individuals from whom the data are collected. Estimates are sufficient to do corrleation

and classify subjects into "high and low categories" for analysis.

Reducing an instrument to a smaller set of items reduces the degree of variation that will be observed on the variable in question. This reduction in the variance in the data increases the beta error (the probability of rejecting a true hypothesis) in the design making both the design and analysis done more conservative statistically; namely, effects related to the "reduced" (or truncated) variable will have to be more pronounced to be significant. The third point that needs to be noted is that the technique of random sampling items from instruments was explored as a potential instrument reducing technique in the present study as this strategy is a strategy that may be employed by researcher without having any empirical data on the instrument in question.

The major weakness in the random sample instrument reduction strategy and procedure is that it assumes that there is no underlying factor structure to the instrument in question which makes some subset of items more far more important that others in generating total scores that are high in construct validity because the items are heavily loaded on the central and key factors the instrument measures. Given this point, a short form of the instrument that was constructed by selecting those items that had the highest factor loadings on the principal factors that accounted for the most variance on the instrument should have the highest construct validity and correlation to total

scores on the full instrument as well as the reliablity. A short form of an instrument that was generated by this procedure should in theory more greatly preserve all of the pyschometric properties in the full instrument including the correlations with other variables in studies of the instrument.

The full 48 item version of the CSDTC was administered to 356 sixth and eighth graders in a middle school in the greater Boston area along with several other instruments. The reason that sixth and eighth graders were used is that marked linear grade level and gender differences were observed by Crandall for these grades (Crandall, 1975).

Four random samples of 12 items (25%) were drawn with replacement from this sample of data to assess the degree to which the random samples correlated with total instrument score. The 48 CSDTC items were factor analyzed using principal component analsysis with communalities in the diagonals and an eigen cut-off value of 1.00 to assess the the factor structure of the CSDTC. A factor analysis of this instrument had not been previously done to the best of this writers knowledge. Three clear factors were found which accounted for 62% of the total variance which is an excellent result given that the response format of the CSDTC is a forced choice true-false format.

The four items that loaded most highly on each of the three major factors found were selected to make Short Form A of the CSDTC (12 items). Short Form B was constructed by

taking the next three most highly loaded item from each factor and then the item that load highest on each of the three factors identified (12 items). Three (3) items on Short Forms A and B of the CSDTC are, therefore, the same. The reasons that this design was used was (1) 8 to 9 items are typically considered to be the minimum for a subtest (Nunnally, 1967), and (2) the overlap gives these two forms a common core that can be used to calibrate the forms and data gathered using the forms on more than one occasion. This overlapping core is a common practice with multiple form instruments (Cronbach, 1990). Given the structure of the parallell short forms of the CSDTC created, one may use the 9 item versions or the 12 item version on any given occasion depending on one's research needs.

The degree to which the the parallell short forms of the CSDTC correlated to each other and long form scores were analyzed as well as the grade level and gender differences observed with each short form. Given that the factor analytical approach was superior to the random samples approach on the aforementioned criteria, a cross-validating replication was done to ensure that the results were not an artifact on the original data set. One hundred and eight one (181) sixth and eighth graders in a different greater Boston community took Form A of the CSDTC, then form B, and then the long form of the CSDTC. These students completed each of the three forms a week apart.

Results

Tables 1 presents the 12 Short Form A items of the CSDTC along with the number of each item on the original scale, how the item is scored, whether or not the item is on both forms, and the factor classification of the item and factor loading of the item in the factor analysis that was done and reported above. Table 2 presents the 12 Short Form B Items of the CSDTC along with the same informations for each item as is given for each Short Form A item in Table 1. As can be seen from Tables 1 and two the factor loading for all items on Short Forms A and B ranged from .63 to .86. The Cronbach alpha coefficient for Short Form A was r=+.73. The alpha coefficient for Short Form B was r=+.72. Both of these alpha coefficient are excellent for 12 dichotomously scored items. The alpha coefficient for all 48 items on the CSDTC was r=+.85. Another point that may be noted from Tables 1 and 2 is that the 12 CSDTC items for each form are arranged in an order that yield no true-false response pattern nor item clusters by factors. Items were arranged in the order given in Tables 1 and 2 to minimize the response bias effects due to the ordering of the items.

Table 3 presents the intercorrelations between grade, gender, Form A, Form B and all items of the CSDTC for the 181 sixth and eighth grade students from whom data were collected

in the cross-validation study done. As can be seen from Table 3, Form A and B correlated with the full length version of the CSDTC at r=+.89 and r=+.92 respectively. Forms A and B intercorrelated with each other at r=+.84. Homogeneous significant differences between grade levels were found for all forms at the same level as Crandall found in his norming data (r=+.28). The tendency to give the type of socially desirable responses on the CSDTC decreases with age. The expected significant difference in responses by gender was found only on From A (r=+.16). This finding might be due to differences between samples and samples over time (Crandall's data is from the seventies), or to Form A being slightly more sensitive to this factor than Form B due to the manner in which the forms were constructed. No significant diferences were found between mean response level by form, grade level or gender when a 2x2x2 three way repeated measures anova was done. The Cronbach alpha coefficient for Short Form A was r=+.77, the alpha coefficient for Short Form B was r=+.70, and the alpha coefficient for all 48 items on the CSDTC was r=+.87. All of these findings were the same the same as those observed in the short forms development (first) sample of 356 students (see Table 4).

Table 5 presents the intercorrelations the four 25% random samples of CSDTC items and all CSDTC items the short forms development (first) sample of 356 students. As can be seen from Table 5, these intercorrleations range from r=+.59 to r=+.71. Although quite good intercorrelations for 12 item

random samples, all of the intercorrleations in Tables 5 are significantly lower than the intercorrleations between Short Forms A, B, and all CSDTC items given in Tables 3 and 4. As can be seen from Table 5, shortening instruments through randomly sampling procedures is an effective reduction strategy when empirical data is not directly available to the researcher on the instrument, but the factor analytical technique outline in this article are significant superior to randomly sampling technique when large sample empirical data is directly available on the instrument in question, as can be seen from the correlations presented in Tables 3 and 4.

A test-retest reliability study was done of Form A and Form B. One seventh grade class (N=26) in a school different from the others in this study was given Form A and another seventh grade class (N=29) was given Form B. Four days later, each class was readiminstered their original form. The test-retest reliability coefficient for Form A was r=+.87 and for Form B was r=+.83. These test-retest reliability coefficients are excellent for 12 item scales, particularly as the test-retest reliability coefficient for all 48 CSCTC items is r=+.89. This finding, it should be noted, also has implications relative to current psychometric theory about the relationship between test length and test reliability. The current prevailing belief that longer tests of a given variable are inherently more reliable than shorter tests was establish before and outside of a factor analytical framework. Short tests can be as reliable as long tests if

the items in the short test are "highly loaded" measures in the factor analytical sense of this term, as the results of the present study show.

Lastly, in order to show that the short forms of the CSDTC developed are actually empirically beneficial in the analysis of real reseach data, 254 seventh, eighth, and nineth grade students were administered Short A of the CSDTC and then a 20 item true-false tests of beliefs in myths about acohol and drinking. The myth belief scale was "key scored" in terms of correct answered and then the total number of myths each student believed in was converted to a percentage score. The median social desirability score for the sample was used to divide students into high and low social desirability classifications rather than the median for each grade to divide that grade level. The median score for the whole sample was used to keep the metric onstant in the analsysis.

Table 6 presents the mean alcohol myth belief percentage levels by grade, gender, and tendency to give socially desirable responses (N=254). As can be seen from Table 6, significant main effects for grade and social desirability levels were found as well as significant grade by gender and gender by social desirability interactions. As can be seen from Table 6, the tendency to belief in alcohol myths decreased with age for those students who had a low tendencies to give social desirable responses, whereas myth belief level tended not to decrease with age for those

students who had a high tendencies to give social desirable responses.

As can be seen in Table 6, the tendency of middle school students to give socially desirable response is significantly correlated to their tendency to agree with myths about alcohol and drinking. The tendency of students to give socially desirable response, therefore, cloud the analyses and interpretation of this type of data. The availablity of a short form of the CSDTC allows this intervening variable to be measured and controlled in the analysis and evaluation of the data. The availablity of parallel short forms of the CSDTC, therefore, has great practical value to researchers who do research on sensitive and/or controversial topics such as alcohol and drugs by questionnaire or interview techniques, as can be seen from the results present in Table 6.

Discussion

The results of the present study empirically show that long instruments can be effectively reduced for research purposes in a manner than will preserve their psychometric properties and construct validity by either random sampling or factor analytical techniques. The present study, however, also showed empirically that factor analytical instrument reduction strategies and techniques are superior to random

sampling strategies and techniques and should be used when the appropriate data is available. These findings, therefore, should help researcher and program evaluators to improve the complexity, sophistication and quality of their designs and studies by employing reduced versions of measurement instrument when and where appropriate.

The present study developed two parallel short forms of the Crandall Social Desirability Scale for Children. These two parallel short forms should be of great use and utility for a variety of researchers, but particularly those who do research on sensitive and/or controversial topics such as alcohol and drug usage by questionnaire or interview techniques. The two short forms developed have great practical value.

The present study also showed that the short forms of the CSDTC developed were empirically helpful in analyzing data. The tendency of middle school students to give socially desirable response was significantly correlated to their tendency to agree with myths about alcohol and drinking. The tendency of students to give socially desirable responses, therefore, clouded the analyses and interpretation of the data, and the availablity of the short form of the CSDTC allowed this intervenening variable to be measured and controlled in the analysis and evaluation of the data.

Lastly, professionals who develop instruments (but particularly long instruments) should use their confirmatory data to develop parallel short forms of their instruments

that may be used by researchers. The developement of empirical short forms for instruments should become standard practice. As researchers gather data with longer instruments, they should develop short forms and send the short forms developed to Buros or ERIC or both sources for cataloguing and use by other researchers. As can be seen from the results of the present study, instrument length in not fixed in concrete and instruments may be reduced in length without reducing or altering their psychometric properties or construct validity. This finding should help to improve the complexity, sophistication and quality of reserach in many areas. The results of this study need to be replicated with other kinds and types of instruments.

References

Crandall, V. (1975). A children's social desirability questionnaire for chilren. Journal of Consulting Psychology, vol. 29, no. 1, February, 28-36.

Cronbach, L. (1951) Coefficient alpha and the internal structure of tests. Psychometrica, 16, 297-334.

Cronbach, L. (1990). Essentials of Psychological Testing. Fifth edition. Harper and Row, New York.

Nunnally, J. (1967). Psychometric Theory. McGraw-Hill, New York.

Table 1: Short-Form A Items of the Crandell Social Desirability Test for Children.

KEY: ( )= Number of item on original scale
[ ]= Correct answer for item; T=true, F=false
*= Item on both forms
{ }= Factor loading of item
F= Factor classification of item (1,2,or 3)

SHORT FORM A (TRUE/FALSE RESPONSE FORMAT)

1. I am always respectful to older people. (5) [F] {.77}F1
2. Sometimes I do not feel like doing what my teachers want me to do. (7) [T] * {.86}F2
3. Sometimes I have felt like throwing things or breaking them. (13) [T] {.83}F1
4. I never act "fresh" or "talk back" to my parents. (8) [F] {.77}F2
5. When I make a mistake, I always admit that I am wrong. (9) [F] {.76}F2
6. I sometimes feel like making fun of other people. (39) [T] {.75}F3
7. I always wash my hands before every meal. (31) [F] {.78}F2
8. Sometimes I wish I could just "mess around" instead of having to go to school. (30) [T] {.77}F3
9. I have never been tempted to break a rule or a law. (34) [F] * {.83}F3
10. Sometimes I dislike helping my parents even though I know they need my help around the house. (32) [T] {.77}F1
11. Sometimes I say things just to impress my friends. (15) [T] * {.84}F1
12. I never shout when I feel angry. (17) [F] {.81}F3

Table 2: Short-Form B Items of the Crandell Social Desirability Test for Children.

KEY: ()= Number of item on original scale
[]= Correct answer for item; T=true, F=false
*= Item on both forms
{}= factor loading of item
F= Factor classification of item (1,2,or 3)

SHORT FORM B (TRUE/FALSE RESPONSE FORMAT)

1. I sometimes feel angry when I don't get my way. (36) [T] {.72}F2
2. I never say anything that would make a person feel bad. (22) [F] {.73}F1
3. Sometimes I argue with my mother to do something she doesn't want me to. (21) [T] {.66}F1
4. I have never been tempted to break a rule or a law. (34) [F] * {.83}F3
5. I would never hit anyone who was smaller than I am. (6) [F] {.66}F3
6. Sometimes I do not feel like doing what my teachers want me to do. (7) [T] * {.86}F2
7. I am always polite even to people who are not very nice. (24) {.73}F2
8. I never borrow anything without asking permission first. (40) [F] {.71}F3
9. Sometimes I say things just to impress my friends.(15) [T] * {.84}F1
10. I am always careful about keeping my cloths neat and my room picked up. (16) [F] {.63}F3
11. Sometimes I don't feel like obeying my parents. (46) [T] {.73}F1
12. Sometimes I feel like staying home from school even if I am not sick. (18) [T] {.71}F2

Table 3: Intercorrelations of Grade, Gender, Form A, Form B, and All Items of Crandall's Social Desirability Scale for Children (N=181).

	GR	GN	FA	FB	ALL
Grade (GR)	1.00	.09	-.29*	-.24*	-.27*
Gender (GN)		1.00	-.16*	-.08	-.10
Form A (FA)			1.00	.84*	.89*
Form B (FB)				1.00	.92*
All (ALL) CSDTC items					1.00

*=p<.05

Table 4: Intercorrelations of Grade, Gender, Form A, Form B, and All Items of Crandall's Social Desirability Scale for Children for the Short Form Development Sample (N=356).

	GR	GN	FA	FB	ALL
Grade (GR)	1.00	.03	-.28*	-.26*	-.28*
Gender (GN)		1.00	-.13*	-.04	-.08
Form A (FA)			1.00	.85*	.88*
Form B (FB)				1.00	.90*
All (ALL) CSDTC items					1.00

*=p<.01

Table 5: Intercorrelations of the Four 25% random samples of Items and All Items of Crandall's Social Desirability Scale for Children (N=356).

	RS1	RS2	RS3	RS4	ALL
Sample 1 (RS1)	1.00	.68	.66	.61	.71
Sample 2 (RS2)		1.00	.59	.67	.68
Sample 3 (RS3)			1.00	.67	.65
Sample 4 (RS4)				1.00	.70
All (ALL) CSDTC items					1.00

All correlations are significant

Table 6: Mean alcohol myth belief percentage levels by grade, gender and tendency to give socially desirable responses (N=254).

	Males				Females			
	Low SD		High SD		Low SD		High SD	
Grade	n	Mean	n	Mean	n	Mean	n	Mean
7	16	33.8	17	31.2	15	30.0	32	28.7
8	16	9.4	17	24.1	25	23.6	32	26.6
9	22	12.2	19	25.6	22	9.5	19	17.7

Source	df	Mean Sq.	F	p
Grade (GR)	2	4330.3	13.2	<.001***
Gender (GN)	1	2.5	0.0	>.05
Soc.Des.(SD)	1	2000.8	6.1	<.01**
GR x GN	2	1129.5	3.4	<.05*
GR x SD	2	897.8	2.7	<.07*
GN x SD	1	368.6	1.1	>.05
GR x GN x SD	2	215.4	0.6	>.05
Error	243	328.9		