Published in final edited form as:

Clin Gerontol. 2009 October 1; 32(4): 371–378. doi:10.1080/07317110903110385.

Does “no” mean “sometimes”? How older adults respond to the same depression symptoms with different response formats

Emily Schoenhofen Sharp,

University of Southern California, Department of Psychology, 3620 McClintock Ave, Los Angeles, CA, 90089. Phone: 213-740-7555, Fax: 213-746-5994, schoenho@usc.edu

Kristen M. Suthers,

American Public Health Association, 800 I Street NW, Washington, D.C., 2001, Phone: 202-777-2434, Fax: 202-777-2532, Kristensuthers@yahoo.com

Eileen Crimmins, and

University of Southern California, Andrus Gerontology Center, 3715 McClintock Ave, Los Angeles, CA 90089, Phone: 213-740-5156, Fax: 213-740-0792, Crimmin@usc.edu

Margaret Gatz

University of Southern California, Department of Psychology, 3620 McClintock Ave, Los Angeles, CA, 90089. Phone: 213-740-2212, Fax: 213-746-5994, gatz@usc.edu

Abstract

In this report we describe problems associated with the administration of binary choice response questionnaires, with particular attention to depression measures given to older adults. A convenience sample of 77 respondents aged 70+ completed two different versions of the 8-item Center for Epidemiologic Studies-Depression (CES-D) scale. Versions were identical except for having either two- or four- response option formats. Within-person responses were compared to determine equivalence across formats. We found that a binary-response option format over- or under-estimated depressive symptomatology. Thus, a four-response option for the CES-D may be a more precise estimate of currently experienced symptoms.

Keywords

response format; survey methods; depression; older adults

The purpose of this study was to (1) describe concerns about the use of binary response option formats, and (2) conduct a direct comparison of two versions of the CES-D, a binary choice yes/no option and a four-choice response option, in a sample of older adults.

The Center for Epidemiologic Studies-Depression (CES-D) scale is widely used in population studies, to estimate the level of depressive symptoms in a given sample, and clinically as a screening instrument for depression. The original CES-D consists of a twenty-item questionnaire with a four-response option format (Radloff, 1977). Numerous variations of the original scale have been used in clinical as well as in field studies of older adults (Kohout, Berkman, Evans, & Cornoni-Huntley, 1993; Turvey, Wallace, & Herzog, 1999). One change has been to include fewer symptoms with the goal of reducing respondent burden and

improving retention rates in longitudinal studies (Kohout et al., 1993; Steffick, 2000; Turvey et al., 1999). Another change has been to reduce the four-response format to three or two response options, with the stated purpose of making the instrument less cognitively demanding for older adults as well as easier to administer verbally (Turvey et al., 1999).

Based on the rationale of administrative ease and reduced cognitive burden for older adults, the authors of the Geriatric Depression Survey (GDS) adopted a binary response format (Yesavage, 1983). However, when using the GDS as a screening instrument for older adults, Fischer, Rolnick, Jackson, Garrard and Luepke (1996) discovered that over a fifth of the surveys (890 out of 3870) included extraneous comments and notations. Examining this response pattern more closely, they found that two-fifths of the total comments were criticizing the yes/no format as too restrictive. Contrary to the hypothesized lessening in frustration, these observations suggest that the binary response option format may actually increase both cognitive demand and participant burden.

A further concern about changing response formats of an established scale is that scores become non-comparable. Prevalence estimates of depressive symptoms in samples of older adults based on the CES-D are dependent upon the adaptation of the instrument that was used. However, when response options are changed, investigators are left to contemplate what a particular score on one type of CES-D response format indicates on another response format. If scores prove to differ with different response formats it would have implications for using cut off scores to estimate prevalence or for clinical screening.

Investigators affiliated with the Health and Retirement Study used a four-response option version of an 11-item CES-D (adopted from the Established Populations for the Epidemiological Studies of the Elderly) in their first wave of data collection (Steffick, 2000). In all subsequent waves, an 8-item yes/no response option CES-D was included in the questionnaire. It was concluded that there was sufficient disagreement across items in the two versions to make it difficult to accurately compare the two CES-D versions. In particular, designating “some of the time” as equivalent to a “yes” response overstated the endorsement of symptoms, while designating “some of the time” as equivalent to “no” understated the endorsement of symptoms (Steffick, 2000).

A more thorough attempt at solving the HRS CESD-D dilemma was undertaken by Jones and Fonda (2004). These authors utilized item response theory (IRT) (see Embretson & Reise, 2000) to create a measurement model. However, their purpose was not to provide a discussion of the clinical implications.

To address these concerns, we conducted a direct comparison of the CES-D using a binary-response and four-response option formats.

Methods

Sample

A convenience sample of 83 older adults was drawn from local communities in Southern California; 6 of these respondents were eliminated from the sample because they failed to answer one or more items on the CES-D scales. Participants were recruited from local churches, independent living retirement communities, a low-income housing complex for seniors, as well as staff and volunteers from the University of Southern California. All participants were living independently in the community at the time of interview. Of the 77 respondents included in these analyses, ages ranged from 60 to 86 years, with a mean age of 72.4 (S.D. = 7.3). The sample was 61.0% female and 39.0% male. The ethnic breakdown was as follows: white=75.3%, black=11.7%, Asian or Pacific Islander=10.4%, unidentified=2.6%.

Educational attainment was treated as a continuous variable ranging from 1 (some high school) to 5 (graduate school). Education was not correlated with either version of the CES-D and did not serve as a moderator of any reported results.

Measures

We administered an eight-item CES-D, with question items corresponding to the version used in the nationally representative Assets and Health Dynamics among the Oldest Old and the second wave of data collection in the HRS (Steffick, 2000; Turvey et al., 1999). These eight items included six negatively valenced items (e.g., feeling sad) and two items that are positively worded (e.g., feeling happy). Respondents were given two versions of the 8-item CES-D scale, one with a four-response option format and one with a binary response format. The four-response format asked respondents to choose either “not at all / almost not at all”, “seldom”, “often”, or “all the time / nearly all the time”, scored 0, 1, 2, or 3, as in the original CES-D with 24 as the maximum possible score. The binary format was a simple “yes/no” option, with each item scored 0 or 1, with 8 as the maximum possible score. In calculating scores, positively worded items were reverse coded. A distracter task of a simple maze was presented between administrations of different versions, with order of administration counterbalanced across participants. There were no significant differences according to order of scale administration.

Results

The mean for the total CES-D scores based on the binary response format was 1.6 (S.D.=2.0) and for the four-response format was 6.1 (S.D.=4.6). When presented with negatively valenced CES-D items on the four response format, respondents tended to endorse at least some feeling of depressed mood. However, when presented with a yes/no choice on negatively valenced items, respondents were less likely to endorse symptoms of depressed mood. As evident in Table 1, for negatively valenced items, a substantial proportion of individuals who endorsed “no” on the binary format chose “seldom” when given four options, but some who endorsed “yes” on the binary format also chose “seldom”.

A greater dispersion of responses was observed on the two positively valenced, positively worded, items. For instance, someone who responded “yes” on the two-response option item, “I was happy”, might have given any of the answers for the four-response option format, including five individuals who responded “not at all / almost not at all”. It is possible that these five individuals may have missed the fact that some of the CES-D items were worded positively. Overall, the spread of the responses suggests that a binary choice format on positive CES-D items may miss more nuanced experiences of mood.

Two ordinary least squares (OLS) regression models were computed to determine a conversion estimate for response formats, using the formula $\hat{\alpha}_j = \beta_{xj} + \alpha$. Each of the response formats was specified in an OLS model as either the dependent or independent variable, respectively. If x is the score on the four-response option measure, then the formula would calculate the equivalent binary-option score ($\hat{\alpha}$). The regression formula derived for converting a four-response option total score to a two-response option total score, was $\hat{\alpha}_j = 0.29 (x_j) - 0.15$, and had confidence intervals for the parameter estimate of β between 0.26 and 0.30. The regression formula derived for converting a two-response option total score to four-response option total score was $\hat{\alpha}_j = 3.42 (x_j) + 1.63$ and had confidence intervals for the parameter estimate of β between 1.43 and 1.83. These two formulae do not provide an exact reciprocal conversion score due to the error associated with beta. However, what is clear is that individuals scoring 0 or 1 on the two-item response option version had scores ranging from 0 to 11 on the four-item version. In turn, those scoring 6 and higher on the two item response option had scores from 8 to 23 on the four-item version. In other words, the binary response format under-

estimates mild symptoms at the low end and over-estimates symptoms at the high end of the scale.

Discussion

Our goal in this report was to (1) describe concerns over the precision of the binary response option format, and (2) compare responses between questionnaires with a four-response option format or a two-response option format.

Previous writers have suggested that the yes/no binary format produces the least response burden. In contrast, other field and clinical research has highlighted pitfalls of a binary response format (see Fischer et al., 1996). In our own experience with the administration of binary format depression measures in both clinical and research settings, we have observed that older adults have difficulty choosing between “yes” and “no”. We have witnessed clients altering the questions by inserting “sometimes”, writing comments in the margin, and either crossing out or inserting words. When verbally administered, we have had to re-direct respondents to choose either “yes” or “no”. In turn, it becomes difficult for the interviewer or data coder to interpret answers that have been modified or that are only half-heartedly endorsed. We suggest, therefore, based on experience and reported research results (Fischer et al., 1996), that the binary response option format may actually create an increased response burden and cognitive demand because the individual must decide whether an experience was sufficient to warrant completely endorsing the symptom.

Furthermore, we are concerned that endorsing “yes” on the binary response format may be more stigmatizing to older adults in a way that endorsing “most of the time” or “often” may not be. If, as has been suggested, older adults are reluctant to volunteer depressive symptoms (e.g. Lyness, Cox, Curry, Conwed, King, & Caine, 1995), then allowing a more nuanced option may make it easier for the respondent. Perceived stigma related to mental illness and receiving psychological help has been suggested to be of particular relevance to older adults and may lead to obstacles in seeking treatment (e.g. Sirey, Bruce, Alexopoulos, Perlick, Raue, et al., 2001).

Our analyses, using intra-individual scores from different CES-D response formats administered at the same time in counterbalanced order, demonstrate that responses to CES-D items do not directly translate across response formats. For negatively valenced items, “seldom” on the four-response format cannot be assumed to consistently correspond either to “no” or to “yes” on the binary format. For positively valenced items, a “yes” on the binary format cannot be assumed to consistently correspond to either “all the time” or “often” on the four-response format. These findings corroborate what Steffick and colleagues (2000) and Jones and Fonda (2004) determined in attempting to convert scores between the two CES-D versions.

A limitation of this study is that the sample size precluded a more sophisticated item-analysis and the current estimation leads to large standard errors in the regression formulae. It is important to note that these analyses are based on respondent characteristics of a convenience sample of older adults in Southern California. Therefore, we cannot rule out that the regression formulae might change if the same study were conducted with a different sample of older adults.

In conclusion, we recommend that both researchers and clinical interviewers adopt measures with four-response options when evaluating depression in older adults. For future research, we believe it is good practice for both research and clinical investigators to document alternative responses from older adults (e.g. written in or verbal comments, criticisms, etc). These data

may be able to be used in future research to better understand response patterns and cognitive burden on measures of depression in older adults.

Acknowledgments

This research was supported in part by the National Institute of Aging (T32 AG00037, P30 AG17265). The authors thank Larry Thompson for his insightful input to the manuscript. The authors also gratefully acknowledge assistance in data collection from Elizabeth Kunda, Kuro Nagasaka, and Adam Turner.

References

- Embretson, SE.; Reise, SP. Item response theory for psychologists. Mahwah, NJ: Erlbaum; 2000.
- Fischer LR, Rolnick SJ, Jackson J, Garrard J, Luepke L. The Geriatric Depression Scale: A content analysis of respondent comments. *Journal of Mental Health and Aging* 1996;2:125–135.
- Jones RN, Fonda SJ. Use of an IRT-based latent variable model to link different forms of the CES-D from the Health and Retirement Study. *Social Psychiatry and Psychiatric Epidemiology* 2004;39:828–835. [PubMed: 15669664]
- Kohout F, Berkman L, Evans D, Cornoni-Huntley J. Two shorter forms of the CES-D: Depression symptoms index. *Journal of Aging and Health* 1993;5:179–193. [PubMed: 10125443]
- Lyness JM, Cox C, Curry J, Conwell Y, King DA, Cain ED. Older age and the underreporting of depressive symptoms. *Journal of American Geriatrics Society* 1995;43:216–221.
- Radloff LS. The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement* 1977;1:385–401.
- Radloff L, Teri L. Use of the Center for Epidemiological Studies-Depression scale with older adults. *Clinical Gerontologist* 1986;5:119–136.
- Sirey JA, Bruce ML, Alexopoulos GS, Perlick DA, Raue P, Friedman SJ, Meyers BS. Perceived stigma as a predictor of treatment discontinuation in young and older outpatients with depression. *American Journal of Psychiatry* 2001;158:479–481. [PubMed: 11229992]
- Steffick, DE. Documentation of affective functioning measures in the health and retirement study. HRS-AHEAD documentation report DR-005. 2000. Retrieved January 20, 2007, from <http://hrsonline.isr.umich.edu>
- Turvey C, Wallace R, Herzog R. A revised CES-D measure of depressive symptoms and a DSM-based measure of major depressive episodes in the elderly. *International Psychogeriatrics* 1999;11:139–148. [PubMed: 11475428]
- Yesavage JA, Brink TL, Rose TL, Lum O, Huang V, Adey M, et al. Development and validation of a geriatric depression screening scale: A preliminary report. *Journal of Psychiatry Research* 1983;17:37–9.

Table 1

Frequencies of individual responses across two CES-D response formats

CES-D Question	Four-Response Option				Two-Response Option	
	“Not at all”	“Seldom”	“Often”	“All the time”	“No”	“Yes”
I felt depressed	53.2% (n=41)	33.8% (n=26)	7.8% (n=6)	5.2% (n=4)	81.8% (n=63)	18.2% (n=14)
Everything was an effort	45.5% (n=35)	37.7% (n=29)	10.4% (n=8)	6.5% (n=5)	76.7% (n=59)	23.4% (n=18)
My sleep was restless	36.4% (n=28)	32.5% (n=25)	26.0% (n=20)	5.2% (n=4)	58.4% (n=45)	41.6% (n=32)
I felt lonely	58.4% (n=45)	29.9% (n=23)	7.8% (n=6)	3.9% (n=3)	79.2% (n=61)	20.8% (n=16)
I was happy*	11.7% (n=9)	9.1% (n=7)	39.0% (n=30)	40.3% (n=31)	14.3% (n=11)	85.7% (n=66)
I felt sad	49.4% (n=38)	37.7% (n=29)	9.1% (n=7)	3.9% (n=3)	81.8% (n=63)	18.2% (n=14)
I enjoyed life*	10.4% (n=8)	5.2% (n=4)	33.8% (n=26)	50.6% (n=39)	11.7% (n=9)	88.3% (n=68)
I could not “get going”	42.9% (n=33)	45.5% (n=35)	5.2% (n=4)	6.5% (n=5)	83.1% (n=64)	16.9% (n=13)

Note.

* denotes items that are positively worded and reverse scored before summing