VU Research Portal

Using Social Desirability Scales in Research among the Elderly

Dijkstra, W.; Smit, J.H.; Comijs, H.C.

document version
Publisher's PDF, also known as Version of record

Link to publication in VU Research Portal

Research Note

Using Social Desirability Scales in Research among the Elderly

WIL DIJKSTRA[1], JOHANNES H. SMIT[2] and HANNIE C. COMIJS[3]
[1]*Department of Social Research Methodology,* [2]*Department of Sociology and Social Gerontology,*
[3]*Institute of Extramural Health Research, Vrije Universiteit, Amsterdam, The Netherlands*

Abstract. A consistent finding is that elderly people obtain higher scores on social desirability (SD) scales than younger ones. It is usually assumed that elderly people are more eager to present themselves in a favourable way. Hence, especially in survey-research among the elderly, it is common practice to include SD-scales to correct for spurious relationships. In this article it is argued that the assumptions underlying such practice are doubtful, because SD-scores may be affected by a variety of factors but the tendency to present oneself favourably. Two such factors related to the question-answering process are discussed in more detail. Empirical support is provided that these factors may explain the relationship between the score on a SD-scale and age. It is argued that using SD-scales to correct for relationships between variables, may lead to erroneous results.

Key words: social desirability scales, data collection, elder respondents, need for approved cognitive abilities

1. Introduction

To measure attitudes, behaviour and opinions in gerontological research, much use is made of standardised questionnaires as method of data collection. Unfortunately, the validity of the data obtained by this method may be hampered by several factors. A well-known disturbing factor is the tendency to present oneself in a more favourable way. If the possible responses on a survey question differ with respect to social desirability, respondents tend to select those alternatives that present themselves in a more favourable way than is actually the case.

It is it is usually assumed that this tendency reflects a more or less stable personality trait like 'need for approval'. This assumption has led to the construction of so-called social desirability scales, or SD-scales for short (Crowne and Marlowe, 1964; Edwards, 1990; Eysenck and Eysenck, 1975; Eysenck and Eysenck, 1990). SD-scales measure the tendency to give socially desirable answers by using questions with response alternatives that differ strongly with respect to their social desirability, the most desirable alternative being highly unlikely to apply

to respondents however. Respondents who select such alternatives apparently are apt to present themselves too favourably. It is assumed that such respondents will also answer other survey questions in a socially desirable direction. Hence, the score on an SD-scale can be used to correct for this tendency. For example, if the scores on two survey questions are correlated and both questions correlate with the score on an SD-scale too, the correlation between both survey question is partly spurious because of the effect of social desirability. In such a case a partial correlation would be a more appropriate indication of the 'true' relation between both survey questions. For this reason, SD-scales are often routinely included in survey-research and added to the theoretical model as a control variable.

SD-scales appear to correlate with age: in general, older respondents obtain higher scores (Eysenck and Eysenck, 1975; Ray, 1988). According to Ray (1988), older people are more prone to present them favourably then younger ones because old age is negatively stereotyped in the western society. Old people therefore – as a kind of defence mechanism – are likely to improve their own self-image by presenting their own virtues in a more favourable way (Shenfield, 1984). Hence, especially in gerontological research it would make sense to use SD-scales to correct for social desirability effects.

However, using SD-scales to adjust for the spurious correlation caused by the tendency to give socially desirable answers, has been criticised. Firstly, much gerontological research concerns topics like life-satisfaction, personality and well-being. SD scales often contain items that are quite similar to the items used in such scales. This raises the question whether correlation between such scales and an SD-scale is a matter of substance or a matter of social desirability (McCrae and Costa, 1983; McCrae, 1986). For example, a correlation between the SD-score and the score on a scale to measure well-being, may be due to the fact that respondents who want to improve their self-image also fake a better well-being (Ray, 1998; Carstensen and Cone, 1986), but also because SD-scales partly measure the respondents' well-being. A correlation between such scales and an SD-scale is hardly surprising then (Kozma and Stones, 1987; McCrae and Costa, 1983). In this way, correction for social desirability may remove true variance from the variable of interest.

Secondly, many items in SD-scales concern statements like 'I am always polite'. Different responses to such items may simply reflect true differences in behaviour, caused by differences in values and norms. Especially elder respondents are known to have more traditional values and norms. Responses on an item like 'I am always polite' thus may reflect differences in values and norms, which in turn may affect the actual behaviour, rather than differences in a tendency to present oneself in a favourable way.

These criticisms in fact concern the validity of SD-scales: do SD-scales measure what they are intended to measure (a tendency to present oneself favourably) or do they measure something else (substantive concepts, differences in norms and values)? A different kind of criticism concerns the question-answering process itself.

In the next section we will present two mechanisms, relevant to this process, that may affect the score on an SD-scale and moreover are related to age. Next we will present empirical data relevant to these issues.

2. Answering SD-items

Many items from an SD-scale may be obvious with respect to their true meaning: to measure if someone is presenting him or herself in a favourable way (Shulman and Silverman, 1974). For example an item like: 'Are *all* your habits good and desirable ones?'. Such an item may suggest the idea that the true purpose is to measure whether or not the respondent speaks the truth. In that case a low score on an SD-scale may indicate a critical attitude towards survey questions, rather than a low tendency to present oneself favourably. It is even possible that respondents who grasp the true intention of SD-scales, tend to select the unfavourable alternatives, to show that they are honest people: favourable and unfavourable alternatives are reversed in a sense. Assuming that such a critical attitude is related to the respondent's educational level, we hypothesise a negative relation between educational level and the SD-score. Because elder respondents usually have had less education, the relation between age and the score on an SD-scale may be partly explained by the intervening variable education, as a proxy for critical attitude.

A second mechanism concerns the respondent's cognitive ability to retrieve the necessary information from memory to answer survey questions. Many items from SD-scales start with 'Have you ever ...?'. For example, 'Have you ever broken or lost something belonging to someone else?'. To answer such a question, the respondent has to retrieve at least one such event from memory. If the respondent simply cannot remember such an event, the most faithful answer is 'No, at least I cannot remember.' Hence, the score on an SD-scale may be determined by the respondent's cognitive ability to retrieve information from memory. The lower this ability, the lower the SD-score. Because this ability tends to decline with age (e.g., Smit et al., 1997; Arbuckle *et al.*, 1992; Birren and Schaie, 1990), older people will have higher scores on an SD-scale than younger people.

The following hypotheses will be tested:
(1) The score on a SD-scale correlates positively with age.
(2) The score on a SD-scale correlates negatively with education.
(3) The score on a SD-Scale correlates negatively with the ability to retrieve information from memory.
(4) The relation between age and the score on a SD-scale diminishes if we control for education and the ability to retrieve information from memory.

3. Method

3.1. RESPONDENTS

The data for this study is collected as part of the Amsterdam Study of The Elderly (Amstel). Amstel is a Dutch community based follow-up study on cognitive functioning and decline in non-institutionalised elderly people (65+ years of age). The first wave in 1990 consisted of 4051 non-institutionalised respondents between 65 and 85 years of age, living in Amsterdam. The second wave took place in 1994, and consisted of 1954 survivors of the original cohort. During the follow-up period a number of side studies were introduced. In 1995 a study investigated cognitive functioning and personality traits as risk factors for elderly abuse among 294 elderly. In this study information on social desirability was also gathered. All respondents were interviewed at their home by five experienced interviewers about topics like aggression and cognitive abilities. More information about the Amstel study and sampling procedures can be found in Launer et al. (1994), more information on the side study on cognitive functioning and elderly abuse can be found in Comijs et al. (1999).

3.2. QUESTIONNAIRE

To measure the ability to retrieve information from memory, three different concepts were used. First a fluency test, a test which is particularly suited to test the functioning of the semantic memory and the ability to retrieve information from memory. The test consisted of mentioning as much names of different animals as possible within one minute. The more names the respondent mentions, the better the ability to retrieve information from memory (Rosen, 1980). Second, a general test to estimate cognitive abilities, the Mini-Mental State Examination (MMSE) (Folstein et al., 1975). The MMSE scores range from 0 (low capacity) to 30 (high capacity). Third, a question from the Cambridge Examination for Mental Disorders of The Elderly (CAMDEX) which measures a subjective opinion of the functioning of the memory of the respondent. 'Do other people say that you are forgetful or confused?', with the answer possibilities: Never (1), Sometimes (2) and Often (3) (Subjective memory) (Roth et al., 1988). The tendency to give socially desirable answers was measured with the Lie-scale from the Eysenck Personality Questionnaire (EPQRS; Eysenck and Eysenck, 1990). This scale consists of 12 items, like 'Have you ever said anything bad or nasty about anyone?'. The respondent can respond with 'yes' or 'no'. The score consists of the number of socially desirable answers. Cronbach's alpha of this test appeared to be 0.66 which can be conceived as sufficient, although not very high.

Education is measured as the number of years of education.

Table I shows the mean, the standard deviation, the minimum and the maximum of these variables.

Table I. Mean, standard deviation, minimum and maximum of the research variables

Variable	Mean	SD	Minimum	Maximum
Age	77.4	5.1	69	89
Fluency	17.1	6.7	3	50
mmse	27.2	2.3	14	30
Subjective memory	1.2	0.4	1	3
Education	8.5	2.4	5	18
SD-score	9.1	2.1	0	12

Table II. Pearson correlations between social desirability scores and age, education and memory

	SD-score	P-value
Age	0.133	0.012
Education	−0.257	0.001
Fluency	−0.194	0.001
mmse	−0.141	0.008
Subjective memory	−0.179	0.001

3.3. ANALYSES

The first three hypotheses were tested by calculating a Pearson correlation and performing a significance test on these correlations on the 5% level. The fourth hypothesis was tested using a structural equation model (LISREL; Jöreskog and Sörbom, 1988).

4. Results

Table II shows the results with respect to the first three hypotheses.

It appears that the score on the social desirability scale correlates positively with age, negatively with education and negatively with cognitive abilities to retrieve information from memory. All correlations are in the expected direction and significant at the 5% level. Hence, hypotheses 1, 2 and 3 are accepted.

Although the correlations are not very high, it is important to note that the relations between the SD-score and education cannot be easily explained in terms of a tendency to present oneself favourably. There is no reason to expect that higher educated people want to present themselves in a less positive way. The negative correlation between the SD-score and the memory variables are even less conceiv-
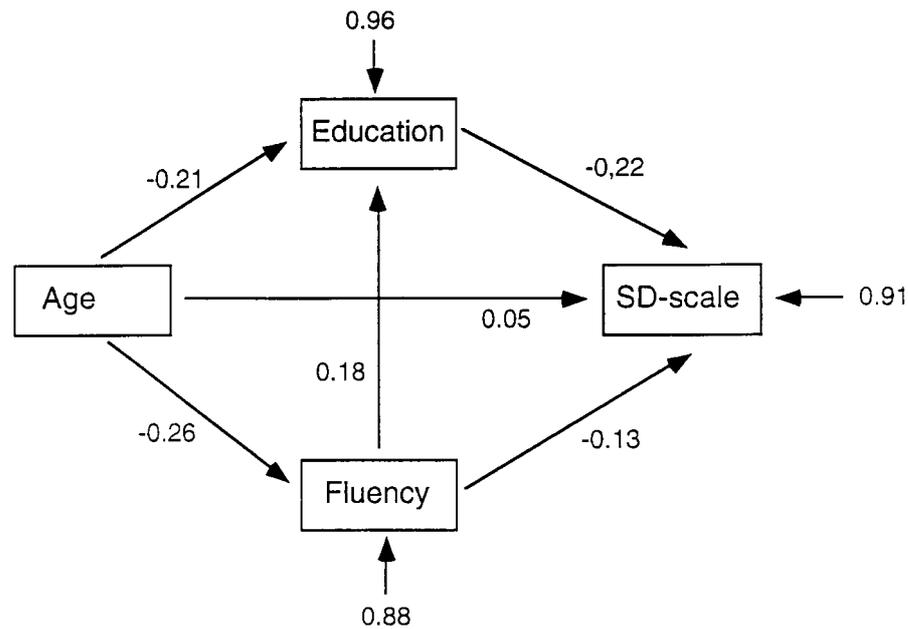
Figure 1. Full model of the effects (*β*'s) of age, education and memory on social desirability score.

able. Hence we conclude that the results support our alternative explanations of determinants of the score on an SD-scale.

To test hypothesis 4, we investigated the consequences for the relation between age and SD when education and memory variables were incorporated in one model. First we specified a saturated model (see Figure 1). For reasons of simplicity we assumed a causal relation from memory to education; the results are essentially the same if the causality is reversed. As memory variable the fluency test was used as an example, as this variable had the highest (negative) correlation with social desirability.

In accordance with hypothesis 4, the effect from age on the SD-score appears to be lower than the correlation between the SD-score and age and is not significant anymore. The other effects are significant at the 5% level. Hence we specified an adjusted model, removing the effect from age on the SD-score (Figure 2).

The data appeared tot fit well into the model. The most important fit measures are: chi-square = 0.70, df = 1, p = 0.40; Adjusted Goodness of Fit Index (AGFI) = 0.99; Root Mean Square Residual (RMR) = 0.014 (Rayof et al., 1991). For the MMSE and subjective memory specific effect estimates differed slightly (results not shown) but had in principle the same result, the effect of age on the SD-scores disappeared. We conclude that the relation between age and SD-score can be fully explained by education and memory. Hypothesis 4 is accepted too.
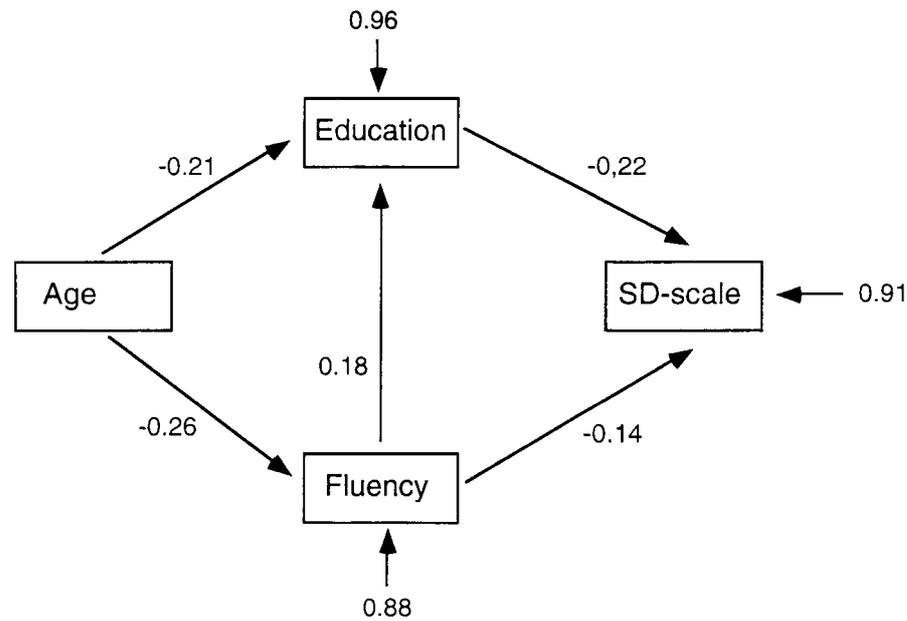
Figure 2. Adjusted model of the effects (β's) of age, education and memory on social desirability score.

5. Discussion

The results suggest that using SD-scales to correct for the tendency to present one-self favourably may lead to incorrect results. In the introduction it was discussed that using such procedures one may effectively correct for substantive variables as measured with personality inventory scales like well-being or life satisfaction, or with different norms and values. In addition we showed that the question-answering process itself also may affect the scores on an SD-scale with the consequence that correction procedures also may partial out variables like cognitive abilities or education, instead of a bias towards a positive self-presentation. Thus, correc-tion procedures may lead to erroneous conclusions about the relation between the variables of interest in a research. The negative correlations between the SD-score on the one hand and education and especially memory on the other hand casts doubt on the validity of SD-scales as instruments to measure a tendency to give socially desirable answers as a stable personality trait. We do not suggest that people don't give socially desirable answers; quite to the contrary, there is ample evidence that people tend to answer questions in a socially desirable direction. We doubt, however, that some people *systematically* tend to answer questions in a socially desirable way, whereas other people *systematically* answer survey questions honestly. Rather, we suspect that social desirability effects are caused by differences in the perception of how socially desirable particular answers are. For example, an item like 'Have you ever taken anything (even a pin or button) that

belonged to someone else?'. Some people may view such behaviour as very undesirable, whereas other people may view such behaviour as not undesirable at all. Hence, if two persons answers this questions differently (with 'yes', respectively 'no') this does not mean that both persons differ with respect to some tendency to present themselves favourably, but with respect to different perceptions of how (un)desirable such behaviour is (Hartmann, 1993; Phillips and Clancy, 1972). Such different perceptions may in turn depend on different norms about behaviour, as already discussed in the introduction.

Our results emphasise the need for further research about the validity of SD-scales. Rather a larger range of age should be used than in the present study, to test hypothesis four. Instead of education, a more direct measure of grasping the intention of SD-scales by respondents would be preferred. Measures about norms (e.g., more or less conservative norms) might give additional insight.

Acknowledgement

References

Arbuckle, T. Y., Gold, D. P., Schwartzman, A. D. & Chaikelson, J. (1992). The role of psychological context, age and intelligence in memory performance in older men. *Psychological Aging* 7: 25–36.

Birren, J. E. & Schaie, K. W. (1990). *Handbook of the Psychology of Aging* (3rd edn). New York: Van Nostrand Reinhold.

Carstensen, L. L. & Cone, J. D., (1983). Social desirability and the measurement of psychological well-being in elderly persons. *Journal of Gerontology* 38: 713–715.

Comijs, H. C., Jonker, C., van Tilburg, W. & Smit, J. H. (1999). Hostility and coping capacities as risk indicators of elder mistreatment. *Social Psychiatry and Psychiatric Epidemiology* 34: 48-52.

Crowne, D. P. & Marlowe, D. (1964). *The Approval Motive*. New York: Wiley.

Edwards, A. L. (1990). Construct validity and social desirability. *American Psychologist* 45: 282–289.

Eysenck, H. J. & Eysenck, S. B. G. (1975). *Manual of the Eysenck Personality Questionnaire (Adult and Junior)*. London: Hodder & Stoughton.

Eysenck, H. J. & Eysenck, S. B. G. (1990). *Manual of the Eysenck Personality Scales (EPS Adult)*. London: Hodder & Stoughton.

Folstein, M. F., Folstein S. E. & McHugh, P. R (1975). Mini-mental state; a practical method for grading the cognitive state of patients for the clinician, *Journal of Psychiatric Research* 12: 189–198.

Hartmann, P. (1993). Controlling response bias by social desirability scales. In: D. Krebs & P. Schmidt (eds), *New Directions in Attitude Measurement*. Berlin: de Gruyter, pp. 338–349..

Jöreskog, K. G. & Sörbom, D. (1988). *Lisrel 7. Guide to Program and Application*. Chicago: Scientific Software International.

Kozma, A. & Stones, M. J. (1987). Social desirability in measures of subjective well-being: A systematic Evaluation. *Journal of Gerontology* 42: 56–59.

Launer, L. J., Wind, A. W. & Deeg, D. J. H. (1994). Nonresponse pattern and bias in a community-based cross-sectional study of cognitive functioning among the elderly. *American Journal of Epidemiology* 139: 803–812.

McCrae, R. R. (1986). Well-being scales do not measure social desirability. *Journal of Gerontology* 41(3): 390–392.

McCrae, R. R. & Costa, P. T. (1983). Social desirability scales: More substance than style. *Journal of Consulting and Clinical Psychology* 51: 882–888.

Phillips, D. L. & Clancy, K. J. (1972). Some effects of 'Social Desirability' in survey studies. *American Journal of Sociology* 77: 921–940.

Ray, J. J. (1988). Lie scales and the elderly. *Journal of Personality and Individual Differences* 9: 417–418.

Rayof, Y., Tomer, A. & Nesselrode, J. R. (1991). Reporting structural equation modelling results in psychology and aging: some proposed guidelines. *Psychological Aging* 6: 499–503.

Rosen, W. G. (1980). Verbal fluency in aging and dementia. *Journal of Clinical Neuropsychology* 2: 135–146.

Roth, M., Huppert, F. A., Tym, E. & Mountjoy C. Q. (1988). CAMDEX. The Cambridge Examination for Mental Disorders of the Elderly. Cambridge: Cambridge University Press.

Shenfield, M. E. (1984). The development course of defense mechanisms in later life. *International Journal of Aging and Human Development* 19: 22–28.

Shulman, A. D. & Silverman, I. (1974). Social desirability and need for approval: Some paradoxical data and a conceptual reevaluation. *British Journal of Social and Clinical Psychology* 13: 27–32.

Smit, J. H., Deeg, D. J. H. & Schmand, B. A. (1997). Asking the age question: a reverse record check study. *Journal of Gerontology* 52: 175–177.