Social desirability does not confound reports of wellbeing or of socio-demographic attributes by older women

SHARRON E. DAWES*, BARTON W. PALMER*†,
MATTHEW A. ALLISON*, THEODORE G. GANIATS*
and DILIP V. JESTE*

ABSTRACT

This study assesses the relationship of social desirability response bias with self-reported physical, mental and cognitive health, successful ageing, and socio-demographic attributes among 1,860 older women at the University of California, San Diego's Clinical Center for the Women's Health Initiative and the Sam and Rose Stein Institute for Research on Aging. The women were aged between 57 and 91 years and lived in the San Diego community. Measures included a ten-item Marlowe–Crowne Social Desirability Scale, and self-report scales of physical, mental and cognitive health, successful ageing and wellbeing, as well as standard socio-demographic attributes. Bivariate correlation and multiple regression models indicated that social desirability scores negatively associated with self-reported levels of hostility, anxiety, perceived stress and self-reported cognitive failures, and that they predicted additional variance in multiple regression analyses above models containing socio-demographic predictors alone. On the other hand, even the strongest associations were what are generally considered 'small effects' ($r < 0.30$). Overall, while the findings support the general validity of most of the self-report measures in studies of normal and successful ageing, consideration of social desirability response bias in the interpretation of self-reports of low levels of some key constructs (anxiety, hostility, stress, self-perceived cognitive deficits) is warranted.

KEY WORDS – social desirability, successful ageing, health, cognition, depression.

Introduction

The continuing rise in the number and proportion of adults aged 65 or more years has made the importance of understanding the process of normal ageing more salient (UNO, 2001). Since the mid-1970s,

* Department of Psychiatry, University of California, San Diego, California, USA.
† San Diego Veterans Medical Research Foundation, San Diego, California, USA.

considerable research attention has focused on neuropsychiatric and other medical disorders prevalent in old age, but efforts to understand and foster optimal wellbeing among older adults also require research on standard and even ideal functioning in later life in terms of healthy physical, cognitive and emotional functioning, as well as overall wellbeing. Researchers that focus on the study of healthy ageing face at least one disadvantage over those focused on illnesses among older adults. Specifically, to study neuropsychiatric or other medical disorders, dense concentrations of potential participants can often be contacted through clinical settings, but studies of broader non-patient populations require recruitment through what are often less concentrated sources. Because studies of healthy ageing may require large, geographically dispersed, non-patient samples, there tends to be an even heavier reliance (than in clinical research) on surveys and other self-report data.

Although self-report data have many advantages over other assessment methods, their validity rests, in part, upon the degree to which the participants give accurate and candid responses. It is important to consider the degree to which responses reflect variance in the targeted construct (the actual state of affairs), as against a socially-desirable response bias (whether or not the latter includes self-deception or intentional impression management) (Edwards 1953; McCrae and Costa 1983). The routine inclusion of measures that directly assess respondents' proclivity toward socially-desirable responses may be of value in clarifying the appropriate interpretation of other self-report data in research on ageing (McCrae and Costa 1983). On the other hand, there is also a danger of over-interpreting responses on social desirability scales, particularly in the absence of information about what other factors may result in responses elevated in a socially-desirable direction.

Consideration of standard models of socially-desirable response biases, as well as the assessment of such biases, illustrates why social desirability may be a particular issue in research on normal ageing. An individual's response on a self-report item seems likely to fall into one of three categories: (a) a valid/accurate reflection of his/her objective state on a given dimension, (b) a valid reflection of his/her own perceptions of his/her state, but (if his/her perception is biased) not an accurate reflection of his/her objective state, or (c) an intentional effort 'to look good' (or, in the case of malingering, 'to look bad'). Researchers frequently have no access to objective information about the actual state of affairs, so they can only draw indirect inferences about which of these categories a participant's response belongs to. A common method of drawing such inferences is to include scales of *social desirability response bias* that have items reflecting socially-valued states that do not actually characterise most peoples' behaviour

(such as invariably driving at or under the speed limit). Physical and/or psycho-social changes associated with normal ageing may evoke alterations in perception or self-presentational concerns, thereby affecting the rates of impression management and of responses in socially-desirable directions (Martin, Leary and Rejeski 2000), and hindering the validity or at least the interpretation of self-report data from older people. On the other hand, interpreting elevated scores on social desirability scales can be problematic in some subgroups, such as those who have been socialised either to place particular importance on the judgement or perceptions by others of their 'moral worth', and/or actually to live by relatively pro-social standards. Some data suggest that older adults respond in a more socially-desirable manner because they are more socially responsible and conscientious, or adhere to more traditional norms and values, so not necessarily because they are lying or 'faking being good' (Dijkstra, Smit and Comijs 2001; McCrae and Costa 1983).

Further complicating the interpretation of self-report and potential socially-desirable response biases are the issues of cognitive changes and of cohort differences in levels of education. For instance, Dijkstra, Smit and Comijs (2001) suggested that the association of age with a tendency toward socially-desirable responses may reflect a confound of education and memory functioning. Older people have been shown to respond in a more socially-desirable manner because they, as a cohort, many have lower education and poorer memories, which affect how well people remember situations during which they did socially-undesirable things, and because both influence the socially-desirable response pattern. Consistent with these findings, other researchers have reported significant associations between responses on social desirability scales and various socio-demographic characteristics. Some of the common predictors of socially-desirable responses are older age, female gender, lower education and/or income, and some ethnic backgrounds (Eisenberg and Okun 1996; Gove and Geerken 1977; Klassen, Hornstra and Anderson 1975; Kozma and Stones 1988; Ray 1988; Stöber 2001; Thomsen *et al.* 2005; Welte and Russell 1993).

Prior studies have also shown significant positive associations between social desirability response style and self-reports of life satisfaction, emotional support, happiness and lack of hostility (Breemhaar, Visser and Kleijnen 1990; Helmers *et al.* 1995; Lu and Shih 1997; Ruehlman, Lanyon and Karoly 1999), and positive attitude to one's own ageing and success in ageing (Abraham and Hansson 1995; Mancini and McKeel 1986). Significant negative associations have also been reported between social desirability and psychological distress (Phillips *et al.* 2006; Ruehlman, Lanyon and Karoly 1999).

To the degree that researchers rely upon self-report data to measure normal variance in various relevant dimensions of normal or successful ageing, it is critical to consider the likelihood that self-reports may be confounded by socially-desirable response styles, including the frequency of socially-desirable responses, the magnitude of such response tendencies (how much of the variance in other self-reports of various dimensions can be accounted for by such a response bias), and to identify the participants' characteristics that associate with impression management or socially-desirable responding. Thus, the purpose of the present study was to identify how socially-desirable responses affect scores on self-report measures after accounting for socio-demographic influences among a large sample of older women. Specifically, the present report describes our study of 1,860 older women who completed a ten-item version of the Marlowe–Crowne Social Desirability Scale (the MCSDS-10; Strahan and Gerbasi 1972). We examined the associations between MCSDS-10 scores, socio-demographic characteristics, and several self-report measures of mental/physical health, cognitive functioning, successful ageing and wellbeing. Based on prior studies of social desirability in the general (non-aged) population, we hypothesised that the MCSDS-10 would account for significant additional variance, above that accounted for by socio-demographic attributes alone, on self-reports of medical burden, subjective wellbeing, self-perceived 'successful ageing', psychological distress, and cognitive functioning. In exploratory analyses, we also examined the frequency of responses in the 'socially desirable' direction on each MCSDS-10 item and the correlation between scores on the individual items with the MCSDS-10 total score.

Methods

The participants

The sample was of middle-aged and older women who participated in the Study of Successful Aging sponsored by the University of California, San Diego (UCSD) Stein Institute for Research on Aging. These participants were recruited from the San Diego Clinical Center cohort of the United States National Institute of Health-funded Women's Health Initiative (WHI). The WHI was a nationwide study of 161,808 generally healthy women aged between 50 and 79 years at the time of enrolment. It included a set of clinical trials and an observational study that focused on prevention of chronic disease post-menopause (Women's Health Initiative Study Group 1998). Women were excluded from the clinical trials if they had a history of breast cancer, sickle cell anaemia or kidney failure, and

from the observational study for medical conditions predictive of a survival time of less than three years or for conditions (such as dementia) inconsistent with study participation. There were 6,427 women in the San Diego WHI cohort: 2,163 in the clinical trials and 4,264 in the observational study.

To enrol in the UCSD Study of Successful Aging, each participant was contacted at her final WHI study visit or by mail, and was invited to complete the Stein Institute's Successful Aging Survey, a 250-item self-administered, self-report questionnaire (reviewed and approved by the UCSD Human Research Protection Program). In response, 1,976 (31%) of the participants returned completed questionnaires. Of these, 1,860 provided complete responses to the MCSDS-10 and were English-speaking women and aged over 55 years. They constitute the analysis sample. We have previously reported some of the data from the Successful Aging Survey (Lamond *et al.* 2008; Moore *et al.* 2007), but the present paper is the first to report a comprehensive analysis of the MCSDS-10 scores. Compared to other WHI participants, those in the analysis sample were more likely to be enrolled in the clinical trial (45% *versus* 29% χ^2 (1) = 142.9, $p < 0.001$), to be Caucasian (89% *versus* 75% χ^2 (1) = 113.7, $p < 0.001$), and to have higher education and income (both $p < 0.001$).

Measures and procedures

Socio-demographic attributes. Age, education, ethnicity and income were obtained by self-report as part of the larger WHI study. Income and education were categorised (*i.e.* for income: ⩽$10,000, $10,000–$19,999, $20,000–$34,999, $35,000–$49,999, $50,000–$74,999, $75,000–$99,999, $100,000–$149,999 and $150,000 or more; and for level of education: No school, Grade school 5–8 years, Some high school (9–11 years), High school diploma/GED (General Educational Development), Vocational training school after high school graduation, Some college/associate degree, College graduate/Baccalaureate degree, Some college/professional school after college graduation, Masters degree, Doctoral degree (PhD, MD, JD, etc.), respectively). For purposes of the present analyses, ethnicity was coded dichotomously as Caucasian or non-Caucasian (Latinos were included in the latter group) although when collected for the WHI, additional categories had been collected (Latino, White, African American, Native American, Chinese, Filipino, Japanese, Korean and Other).

Social desirability. The Successful Aging Survey, which was a slightly modified version of the survey questionnaire described in our earlier studies (Lamond *et al.* 2008; Montross *et al.* 2006; Moore *et al.* 2007), included

a previously validated MCSDS-10 (Reynolds 1982; Strahan and Gerbasi 1972). Each MCSDS-10 item consists of a self-descriptive statement to which the participant responds 'true' or 'false', yielding a potential total score that can range from 0 to 10. Five of the items (1, 2, 5, 6 and 7) are reverse coded to allow a score of one point per item checked in the socially-desirable direction. The MCSDS-10 strongly correlates with the full 33-item schedule ($r = 0.85$) (Reynolds 1982).

Self-reported physical and mental health, cognitive function and wellbeing. All measures have been validated in numerous populations (Boulet and Boss 1991; Edwards, Feightner and Goldsmith 1995; Jenkinson, Wright and Coulter 1994; Johnson, Murphy and Dimond 1996; McHorney, Ware and Raczek 1993; Wijndaele *et al.* 2007). The Successful Aging Survey included the Medical Outcomes Survey 36-item Short Form (SF-36; Ware and Sherbourne 1992), the Philadelphia Geriatric Center Morale Scale (PGCM; Lawton 1975), the Center for Epidemiological Studies Depression Scale (CES-D; Radloff 1977), the Anxiety and Hostility sub-scales of the Brief Symptom Inventory (BSI; Derogatis and Melisaratos 1983), and the Perceived Stress Scale (PSS; Cohen, Kamarck and Mermelstein 1983). Higher scores on the PGCM represent more positive attitudes toward ageing, and higher physical and mental health composite scores on the SF-36 reflect greater wellbeing; higher scores on the CES-D, BSI and PSS reflect more psychological distress. As in prior studies (Lamond *et al.* 2008; Moore *et al.* 2007), the participants also rated their degree of successful ageing on a scale from '1' (least successful) to '10' (most successful). No specific definitions of successful ageing were presented as we wanted the participants to choose their own criteria.

Self-perceived cognitive deficits were evaluated with the Cognitive Failures Questionnaire (CFQ) (Broadbent *et al.* 1982), which assesses self-perceptions of problems in various cognitive dimensions such as attention/distractibility, memory and motor functions (Broadbent *et al.* 1982; Wallace, Kass and Stanny 2002). Higher CFQ scores represent greater self-perceived cognitive problems. The participants also completed a modified version of the Cognitive Assessment Screening Test (CAST; Drachman *et al.* 1996), a self-administered test of actual cognitive deficits. Lower CAST scores represent worse cognitive function. This version of the CAST did not include the Orientation questions (Moore *et al.* 2007).

Statistical analyses

Descriptive statistics for all variables including means, medians, standard deviations (SD), ranges, skew and kurtosis were calculated. For each

MCSDS-10 item, we calculated the proportion of participants who endorsed the item in the 'socially desirable' response direction, and Spearman's correlation between each item with the MCSDS-10 total score. A series of regressions were conducted with and without inclusion of the MCSDS-10 to assess the influence of the MCSDS-10 on responses over and above that of the socio-demographic attributes (age, education, income and ethnicity). Relative differences between the R-squared values were then computed (R^2 from the model including MCSDS-10 divided by R^2 from the model *not* including MCSDS-10) to determine the relative influence of the MCSDS-10 on the response for each test. The association between MCSDS-10 scores and ethnicity (Caucasian versus non-Caucasian) was evaluated by a *t*-test; the bivariate associations between MCSDS-10 total scores and each of the other participant characteristics were evaluated with Pearson's correlation, except for education which was evaluated with Spearman's *rho*. Statistical significance was determined at the 5 per cent level ($p < 0.05$; two-tailed). Given the relatively large sample size, it is possible to identify statistically significant associations that may nonetheless have marginal pragmatic magnitude. Thus, we also describe the magnitude of associations with the descriptors for effect size ranges suggested by Cohen (1988): 'small' ($r \geqslant 0.1$ and < 0.3), 'medium' ($r \geqslant 0.3$ and < 0.5), and 'large' ($r \geqslant 0.5$).

Results

Social desirability

The mean MCSDS-10 score was 6.8 (SD $= 1.9$; range $= 0$–10). Endorsement rates for each of the MCSDS-10 items ranged from 36 per cent ('I like to gossip at times') to 94 per cent ('I always try to practise what I preach'), with seven items being endorsed in the 'socially desirable' direction by a majority of the sample (Table 1). Spearman's correlations between MCSDS-10 items and the total score ranged from $r_s = 0.22$ ('I always practise what I preach') to $r_s = 0.54$ ('There have been occasions when I took advantage of someone' ($p < 0.001$ in all cases). Note that consistent with Cohen's (1988) descriptive scheme, the lowest of these correlations was in the small effect size range ($r_s = 0.1$–0.3); of the remaining nine, five were in the 'medium' effect size range ($r_s = 0.3$–0.5), and four were in the 'large' effect size range ($r_s \geqslant 0.5$).

Bivariate correlations of socio-demographic attributes and social desirability

As shown in Table 2, the correlations between the MCSDS-10 scores and age and education were statistically significant ($r_s = 0.18$ and $r_s = -0.11$,

T A B L E 1. *Proportion of individual item responses in the socially desirable direction and their relationship to total score on the ten-item Marlowe–Crowne Social Desirability Scale (MCSDS-10)*

Item	Endorsed in socially-desirable direction (%)	r_s between item and total scores
1. I like to gossip at times. (False)	36	−0.50
2. There have been occasions when I took advantage of someone. (False)	74	−0.54
3. I am always willing to admit it when I make a mistake. (True)	83	0.40
4. I always try to practise what I preach. (True)	94	0.22
5. I sometimes try to get even rather than forgive and forget. (False)	85	−0.40
6. At times I have really insisted on having things my own way. (False)	39	−0.52
7. There have been occasions when I felt like throwing things. (False)	74	−0.48
8. I never resent being asked to return a favour. (True)	80	0.39
9. I have never been irked when people expressed ideas very different from my own. (True)	38	0.52
10. I have never deliberately said something that hurt someone's feelings. (True)	80	0.45

Notes: All correlations were significant at $p < 0.001$ (negative correlations for items coded as 'false', and positive correlations for items coded as 'true'). r_s: Spearman's rank correlation coefficient. Sample size: 1,860.

respectively; both $p < 0.001$); however, the absolute magnitude of both of these correlations was in the 'small effect size' range (Cohen 1988). There was no significant difference in MCSD-10 scores among Caucasian *versus* non-Caucasian participants, *i.e.* mean (and SD) MCSDS-10 total scores were 6.82 (1.92) and 6.99 (1.94), respectively (t (1,853 degrees of freedom) = 1.20, $p = 0.23$).

Association of social desirability on other self-reports of status

As shown in Table 3, the regression models with only the socio-demographic attributes (the 'reduced model') and those with also the MCSDS-10 scores (the 'full model') were completed to assess the extra variance explained by MCSDS-10 on each test score. The magnitude differences in R-squared between the reduced and full models indicated that the MCSDS-10 contributed most in terms of additional explained variance for models in which the dependent variables were measures of psychological distress (BSI Anxiety and Hostility subscales), *i.e.* the full models for these two variables explained more than ten times the amount of variance than did the corresponding reduced models. Addition of the

T A B L E 2. *Descriptive statistics and bivariate associations between socially desirable responses (ten-item Marlowe–Crowne Social Desirability Scale (MCSDS-10) score) and participant socio-demographic attributes*

Participants' characteristics	Mean (SD) or percentage	Association with MCSDS-10 total score[1]
MCSDS-10 score[2]	6.8 (1.9)	–
Socio-demographic attributes:		
Age in years: mean (SD)	72.8 (7.3)	$r(1857) = 0.18$[3]
Education (years or level):		$r(1843) = -0.11$[3]
⩽11 years	2.4%	
High school diploma or General Educational Development credential	10.6%	
Some college or vocational training	43.3%	
Bachelor's degree	10.7%	
Some postgraduate	33.0%	
Ethnic background:		$t(1853) = 1.20$[4]
(Non-Hispanic) Caucasian	87.6%	
Non-Caucasian	12.4%	
Income (annual):		$r(1824) = -0.08$[3]
⩽$10,000	3.6%	
$10,000–$19,999	9.6%	
$20,000–$34,999	20.5%	
$35,000–$49,999	22.0%	
$50,000–$74,999	23.2%	
$75,000–$99,999	9.0%	
$100,000–$149,999	7.1%	
$150,000+	3.0%	
Do not know	2.1%	

Notes: 1. Figures in brackets are degrees of freedom. 2. Possible range of scores: 0–10. 3. With the current sample size, this result is significantly ($p < 0.001$) different from the null hypothesis value ($\rho = 0$), but the magnitude of the association is in the 'small' effect size range (Cohen 1988). 4. $p = 0.23$.

MCSDS-10 into the model also explained at least twice the amount of variance over the reduced model for self-reported mental health functioning (SF-36 Mental Health component), other measures of psychological distress (CES-D and PSS), and self-reported successful ageing and cognitive impairment (CFQ) (Table 3). There were no magnitude differences between models with or without the inclusion of the MCSDS-10 for assessing physical health (SF-36 Physical Health component score), well-being/attitude toward ageing (Philadelphia Geriatric Center Morale Scale), or actual cognitive deficits (CAST) (relative increase of full *versus* reduced models all = 1).

As noted above, the two models in which the addition of the MCSDS-10 accounted for the most additional variance in the dependent variable, over the reduced model alone, were those for the BSI Anxiety and

TABLE 3. *Descriptive statistics and change in R² for predicting participant responses with and without accounting for socially desirable responses (ten-item Marlowe–Crowne Social Desirability Scale (MCSDS-10) score)*

Test	Mean (SD)	Range	R^2 Full[1]	R^2 Reduced[2]	RI	MCSDS-10 p	Correlation MCSDS-10[4]	p
Physical and mental health and wellbeing:								
Medical Outcomes Survey SF-36:								
Physical Health component T-score	43.5 (11.2)	8–68[3]	0.093	0.089	1	0.005	r (1816)=0.01	0.846
Mental Health component T-score	56.7 (7.2)	21–74[3]	0.025	0.008	3	<0.001	r (1816)=0.14	<0.001
Philadelphia Geriatric Center Morale Scale	3.8 (1.3)	0–5[3]	0.048	0.042	1	0.001	r (1761)=0.05	0.060
Center for Epidemiological Studies Depression Scale	6.5 (6.2)	0–38	0.042	0.019	2	<0.001	r (1537)=−0.12	<0.001
Brief Symptom Inventory:								
Anxiety subscale	1.4 (2.3)	0–20	0.034	0.003	11	<0.001	r (1839)=−0.17	<0.001
Hostility subscale	0.9 (1.3)	0–20	0.061	0.001	61	<0.001	r (1840)=−0.24	<0.001
Perceived Stress Scale	9.7 (5.4)	0–32	0.085	0.015	6	<0.001	r (1784)=−0.23	<0.001
Self-rated 'successful ageing'	8.2 (1.4)	1–10[3]	0.024	0.012	2	<0.001	r (1858)=0.09	<0.001
Cognitive deficits or concerns:								
Cognitive Failures Questionnaire total	46.2 (8.7)	0–100	0.109	0.029	4	<0.001	r (1790)=−0.24	<0.001
Cognitive Assessment Screening Test	18.5 (2.2)	0–21[3]	0.088	0.087	1	0.615	r (1794)=−0.06	0.021

Notes: Sample size: 1,860. SD: standard deviation. RI: relative increase. 1. Includes age, income, education, ethnicity and MCSDS-10. 2. Includes age, income, education and ethnicity only. 3. Higher scores indicate better functioning. 4. Figures in brackets are degrees of freedom.

Hostility subscales. However, these were also the two dependent variables for which the reduced model alone accounted for the least variance ($R^2 = 0.003$ and 0.001, respectively). Thus to examine further the association of socially-desirable response scores with other self-reported status variables, we also calculated the bivariate Pearson correlations between MCSDS-10 and each of the ten dependent variables (*see* Table 3, columns 8 and 9). None of these correlations exceeded the 'small' effect size range, *i.e.* $\geqslant 0.3$. However, in terms of statistical significance, higher socially-desirable response style (MCSDS-10 total scores) was associated with better self-reported mental health functioning (SF-36 Mental Health component; $r = 0.14$) and less psychological distress (CES-D $r = -0.19$; BSI Anxiety and Hostility subscales $r_s = -0.17$ and -0.24, and PSS $r = -0.23$), as well as less self-reported cognitive impairment (CFQ; $r = -0.24$) (all $p < 0.001$). There were also statistically significant correlations between the MCSDS-10 and self-reported successful ageing ($r = 0.09$, $p < 0.001$) and actual cognitive deficits (CAST total score, $r = -0.06$, $p = 0.021$), were also statistically significant (all $p < 0.05$) but the magnitude of these correlations was below the cut-off of $r > 0.1$ suggested by Cohen (1988) to define even a 'small effect'. The correlations between MCSDS-10 and physical health (SF-36 Physical Health component score) and wellbeing/attitude toward ageing (Philadelphia Geriatric Center Morale Scale) were not statistically significant (both $p > 0.05$).

Discussion

The findings from seven of the ten regression models tested were consistent with our *a priori* hypotheses, in that in these seven models, social desirability accounted for at least twice the amount of variance as that explained by the socio-demographic attributes alone for mental health, self-perceived cognitive deficits, successful ageing, and wellbeing. The MCSDS-10 did not add any interpretive value in two tests, those of physical health and wellbeing/attitude toward ageing. These regression models indicated that although the MCSDS-10 was a significant predictor, it did not add much additional relative value to the variance explained as the relative R-squared for both physical health and wellbeing/attitude toward ageing regressions were essentially equivalent, with or without the MCSDS-10. Therefore, socially-desirable responding did not appear to have much influence in the interpretation of the questions on these tests by the person taking them. The only regression model for which the addition of the MCSDS-10 did not result in significant additional explained variance was that for actual cognitive deficits with

socio-demographic predictors alone. This is interesting as it was the only test not influenced by the responders' opinion or beliefs – it is more of an objective test.

Although previous research has shown that when scores are corrected for socially-desirable responding, this does not improve the observed validity of the scores, *e.g.* the concurrent validity in terms of association with objective external criteria may even be reduced (Dicken 1963; McCrae and Costa 1983; Wrobel and Shapiro 1999), our findings indicate that adjusting for social desirability on some variables is worthwhile and justifiable considering the extra variance accounted for over and above socio-demographic attributes alone. In particular, considering the bivariate associations, as well as the degree of additional variance accounted for in the multiple regression analyses, it appears that respondents with a higher proclivity to give social desirable responses may also tend to report lower levels of anxiety, hostility, perceived stress and self-reported cognitive deficits. Social desirability thus warrants consideration as a partial explanation of inter-respondent differences in survey studies of these constructs. On the other hand, consideration must also be given to the possibility that such associations reflect overlap in item content rather than response biases. For instance, item 7 on the MCSDS-10, 'There have been occasions when I felt like throwing things', may be less applicable to persons with particularly low levels of stress or hostility, and thus result in a negative correlation between the two scales.

We also found statistically significant although small magnitude influences of some socio-demographic characteristics on degree of social desirability response styles. Age and education were positively correlated with scores on the social desirability measure (MCSDS-10), but the magnitudes of these correlations were only $r=0.18$ and ρ (rho)$=-0.11$, respectively, while income was negatively correlated ($r=-0.08$). For ethnicity, no differences were found between Caucasians and other ethnic groups on mean social desirability response scores. The latter finding contrasts with some reports that have suggested an ethnicity effect (Cappeliez 1989). However, our finding regarding level of education is consistent with earlier investigations in other populations (Cappeliez 1989; Gove and Geerken 1977; Klassen, Hornstra and Anderson 1975). It is to be noted though that because of the rather low response rate to the survey and the corresponding potential presence of selection bias, especially considering that all respondents were women, these low magnitude correlations may not reflect the general older population.

In contrast to the small associations between the MCSDS-10 and socio-demographic characteristics, the correlations between individual MCSDS-10 items and the total score were generally high, with all but one of the

ten values being in the ranges commonly labelled as 'medium' or 'large' effect sizes, *i.e.* correlations above 0.30 and 0.50, respectively (Cohen 1988). This finding speaks to the internal consistency reliability of the MCSDS-10. Thus, the low magnitude of associations between the MCSDS-10 and socio-demographic characteristics did not appear to be attributable to a lack of scale reliability.

There have been suggestions in the empirical literature that older women may be more likely to endorse statements in the 'socially desirable' direction (Eisenberg and Okun 1996; Gove and Geerken 1977; Klassen, Hornstra and Anderson 1975; Kozma and Stones 1988; Ray 1988; Stöber 2001; Thomsen *et al.* 2005; Welte and Russell 1993). In that regard, it might be noted that seven of the ten items on the MCSDS-10 were endorsed in the 'socially desirable' direction by a majority of the present sample. At first glance, such high proportions of participants giving responses that appear socially desirable are consistent with the prior suggestions of gender and age effects on social desirability. The age range in our sample spanned 57–91 years and, as noted above, there was a statistically significant but small effect size correlation between current age and MCSDS-10 scores. However, the analysis had limited ability to evaluate age or gender effects, as we lacked data from younger women and from men by which one might evaluate the degree to which the observed pattern is unique to older women.

There are several other limitations to the present study. All measures were self-report, the sample was restricted to the San Diego cohort of the WHI investigation, and only one-third of the San Diego WHI cohort completed the present survey. Also, we used an abbreviated MCSDS, which although well validated (Reynolds 1982; Strahan and Gerbasi 1972), may not have given an accurate reflection of what is considered to be socially desirable in this population. There may be problems with designating certain responses as 'socially desirable'; for example, the participant might not have understood how to interpret literally a given question (with terms such as 'always' or 'never' which are used in common parlance but are not exact), or that the responses were mediated more by self-deception or impression management, two closely-related concepts (Barrick and Mount 1996). The lack of substantial associations between social desirability scores and self-reports of wellbeing (as measured by the SF-36 Mental Health component and PSS) conflicts with some prior reports for other populations (Breemhaar, Visser and Kleijnen 1990; Lu and Shih 1997; Ruehlman, Lanyon and Karoly 1999). This difference could be related to the item content of the specific scales used to operationalise the key constructs. Kozma and Stones (1987, 1988) suggested that the strength of such associations, when present, reflects common item content in scales

of quality of wellbeing and scales of social desirability response bias. For instance, they suggested that the Edwards Social Desirability Scale (ESDS) (Edwards 1957) has greater content overlap with commonly-used quality of wellbeing scales than did the full-length version of the MCSDS (Crowne and Marlowe 1960), and found that the ESDS had substantially higher correlations with three wellbeing scales relative to the correlations between the MCSDS and these wellbeing scales (Kozma and Stones 1988).

To our knowledge, only one prior study of healthy volunteers has described the specific endorsement rates on the individual items in the MCSDS-10. Specifically, in Reynolds's (1982) study of undergraduate college students, some of the most frequently endorsed MCSDS-10 items were the same as those endorsed in the present sample of older women, *e.g.* 85 per cent of the college students responded in a socially-desirable direction to 'I always try to practise what I preach' (true), 77 per cent to 'I never resent being asked to return a favour' (true), 61 per cent to 'I'm always willing to admit it when I make a mistake' (true), 47 per cent to 'I sometimes try to get even rather than forgive and forget' (false), and 41 per cent to 'I have never been irked when people expressed ideas very different from my own' (true). A part of the difficulty in interpreting the rates at which MCSDS items are endorsed in the 'socially desirable' direction is that a socially-desirable response bias cannot itself be measured directly, but rather must be inferred from responses to items that may reflect self-presentational concerns, *i.e.* the truth status for a particular individual on any particular items is generally unknown. While a person consciously endeavouring to present her/himself in a falsely favourable light will tend to answer 'true' to an item such as 'I always try to practise what I preach', a very conscientious person may also validly/candidly respond 'true' if s/he interprets the item as asking whether s/he attempts to live up to her/his stated values. (Whether one perceives oneself as successful in such attempts is a somewhat different question that is not assessed by the MCSDS items). In future studies, one means of disentangling social desirability response styles from accurate reports of socially-desirable behaviour might be to compare collateral and self-report reports of socially-desirable behaviours.

In conclusion, the reported findings indicate that social desirability as a concept has some explanatory power in regard to the variance in self-reports of successful ageing, or any of the other constructs related to physical and mental health, cognitive function and wellbeing in older women, especially when the question being asked requires more individual interpretation.

Acknowledgements

References

Abraham, J. D. and Hansson, R. O. 1995. Successful aging at work: an applied study of selection, organization, optimization, and compensation through impression management. *Journals of Gerontology: Series B: Psychological Sciences and Social Sciences*, **50B**, 2, 94–103.

Barrick, M. R. and Mount, M. K. 1996. Effects of impression management and self-deception on the predictive validity of personality constructs. *Journal of Applied Psychology*, **81**, 3, 261–72.

Breemhaar, B., Visser, A. P. and Kleijnen, J. G. 1990. Perceptions and behaviour among elderly hospital patients: description and explanation of age differences in satisfaction, knowledge, emotions and behaviour. *Social Science and Medicine*, **31**, 12, 1377–85.

Broadbent, D. E., Cooper, P. F., FitzGerald, P. and Parkes, K. R. 1982. The Cognitive Failures Questionnaire CFQ and its correlates. *British Journal of Clinical Psychology*, **21**, 1, 1–16.

Boulet, J. and Boss, M. W. 1991. Reliability and Validity of the Brief Symptom Inventory. *Psychological Assessment: A journal of Consulting and Clinical Psychology*, **3**, 3, 433–437.

Cappeliez, P. 1989. Social desirability response set and self-report depression inventories in the elderly. *Clinical Gerontologist*, **9**, 2, 45–52.

Cohen, J. 1988. *Statistical Power Analysis for the Behavioral Sciences*. Lawrence Erlbaum, Hillsdale, New Jersey.

Cohen, S., Kamarck, T. and Mermelstein, R. 1983. A global measure of perceived stress. *Journal of Health and Social Behavior*, **24**, 4, 385–96.

Crowne, D. P. and Marlowe, D. 1960. A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology*, **24**, 4, 349–54.

Derogatis, L. R. and Melisaratos, N. 1983. The Brief Symptom Inventory: an introductory report. *Psychological Medicine*, **13**, 3, 595–605.

Dicken, C. 1963. Good impression, social desirability, and acquiescence as suppressor variables. *Educational and Psychological Measurement*, **23**, 4, 699–720.

Dijkstra, W., Smit, J. H. and Comijs, C. H. 2001. Using social desirability scales in research among the elderly. *Quality and Quantity*, **35**, 1, 107–15.

Drachman, D. A., Swearer, J. M., Kane, K., Osgood, D., O'Toole, C. and Moonis, M. 1996. The Cognitive Assessment Screening Test CAST for dementia. *Journal of Geriatric Psychiatry and Neurology*, **9**, 4, 200–8.

Edwards, A. L. 1953. The relationship between the judged desirability of a trait and the probability that the trait will be endorsed. *Journal of Applied Psychology*, **37**, 2, 90–3.

Edwards, A. L. 1957. *The Social Desirability Variable in Personality Assessment and Research*. Henry Holt, New York.

Edwards, M., Feightner, J. and Goldsmith, C. H. 1995. Inter-rater reliability of assessments administered by individuals with and without a background in health care. *Occupational Therapy Journal of Research*, **15**, 2, 103–10.

Eisenberg, N. and Okun, M. A. 1996. The relations of dispositional regulation and emotionality to elders' empathy-related responding and affect while volunteering. *Journal of Personality*, **64**, 1, 157–83.

Gove, W. R. and Geerken, M. R. 1977. Response bias in surveys of mental health: an empirical investigation. *American Journal of Sociology*, **82**, 6, 1289–317.

Helmers, K. F., Krantz, D. S., Merz, C. N., Klein, J., Kop, W. J., Gottdiener, J. S. and Rozanski, A. 1995. Defensive hostility: relationship to multiple markers of cardiac ischemia in patients with coronary disease. *Health Psychology*, **14**, 3, 202–9.

Jenkinson, C., Wright, L. and Coulter, A. 1994. Criterion validity and reliability of the SF-36 in a population sample. *Quality of Life Research*, **3**, 1, 7–12.

Johnson, L. C., Murphy, S. A. and Dimond, M. 1996. Reliability, construct validity, and subscale norms of the Brief Symptom Inventory when administered to bereaved parents. *Journal of Nursing Measurement*, **4**, 2, 117–27.

Klassen, D., Hornstra, R. K. and Anderson, P. B. 1975. Influence of social desirability on symptom and mood reporting in a community survey. *Journal of Consulting and Clinical Psychology*, **43**, 4, 448–52.

Kozma, A. and Stones, M. J. 1987. Social desirability in measures of subjective well-being: a systematic evaluation. *Journal of Gerontology*, **42**, 1, 56–9.

Kozma, A. and Stones, M. J. 1988. Social desirability in measures of subjective well-being: age comparisons. *Social Indicators Research*, **20**, 1, 1–14.

Lamond, A. J., Depp, C. A., Allison, M., Langer, R., Reichstadt, J., Moore, D. J., Golshan, S., Ganiats, T. G. and Jeste, D. V. 2008. Measurement and predictors of resilience among community-dwelling older women. *Journal of Psychiatric Research*, **43**, 2, 148–54.

Lawton, M. P. 1975. The Philadelphia Geriatric Centre Morale Scale: a revision. *Journal of Gerontology*, **30**, 1, 85–89.

Lu, L. and Shih, J. B. 1997. Personality and happiness: is mental health a mediator? *Personality and Individual Differences*, **22**, 2, 249–56.

Mancini, J. A. and McKeel, A. J. 1986. Social desirability and psychological well-being reports in late life: a further inquiry. *Educational and Psychological Measurement*, **46**, 1, 89–94.

Martin, K. A., Leary, M. R. and Rejeski, W. J. 2000. Self-presentational concerns in older adults: implications for health and well-being. *Basic and Applied Social Psychology*, special issue, *The Social Psychology of Aging*, **22**, 3, 169–79.

McCrae, R. R. and Costa, J. P. T. 1983. Social desirability scales: more substance than style. *Journal of Consulting and Clinical Psychology*, **51**, 6, 882–8.

McHorney, C. A., Ware, J. E. Jr and Raczek, A. E. 1993. The MOS 36-Item Short-Form Health Survey (SF-36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. *Medical Care*, **31**, 3, 247–63.

Montross, L. P., Depp, C. A., Daly, J., Reichstadt, J., Golshan, S., Moore, D. J., Sitzer, D. and Jeste, D. V. 2006. Correlates of self-rated successful aging among community-dwelling older adults. *American Journal of Geriatric Psychiatry*, **14**, 1, 43–51.

Moore, D. J., Sitzer, D., Depp, C. A., Montross, L. P., Reichstadt, J., Lebowitz, B. D. and Jeste, D. V. 2007. Self-administered cognitive screening for a study of successful aging

among community-dwelling seniors: a preliminary study. *International Journal of Geriatric Psychiatry*, **22**, 4, 327–31.

Phillips, L. H., Henry, J. D., Hosie, J. A. and Milne, A. B. 2006. Age, anger regulation and well-being. *Aging and Mental Health*, **10**, 3, 250–6.

Radloff, L. S. 1977. The CES-D Scale: a self-report depression scale for research in the general population. *Applied Psychological Measurement*, **1**, 3, 385–401.

Ray, J. J. 1988. Lie scales and the elderly. *Personality and Individual Differences*, **9**, 2, 417–8.

Reynolds, W. M. 1982. Development of reliable and valid short forms of the Marlowe-Crowne Social Desirability Scale. *Journal of Clinical Psychology*, **38**, 1, 119–25.

Ruehlman, L. S., Lanyon, R. I. and Karoly, P. 1999. Development and validation of the multidimensional health profile. Part I: Psychosocial functioning. *Psychological Assessment*, **11**, 2, 166–76.

Strahan, R. and Gerbasi, K. C. 1972. Short, homogeneous versions of the Marlow-Crowne Social Desirability Scale. *Journal of Clinical Psychology*, **28**, 2, 191–3.

Stöber, J. 2001. The Social Desirability Scale-17 SDS-17: convergent validity, discriminant validity, and relationship with age. *European Journal of Psychological Assessment*, **17**, 3, 222–32.

Thomsen, D. K., Mehlsen, M. Y., Viidik, A., Sommerlund, B. and Zachariae, R. 2005. Age and gender differences in negative affect: is there a role for emotion regulation? *Personality and Individual Differences*, **38**, 8, 1935–46.

United Nations Organisation (UNO) 2001. *World Population Ageing: 1950–2050.* Population Division, Department of Economic and Social Affairs, UNO, New York. Available online at http://www.un.org/esa/population/publications/worldageing19502050/ [Accessed 01 November 2010].

Wallace, J. C., Kass, S. J. and Stanny, C. J. 2002. The cognitive failures questionnaire revisited: dimensions and correlates. *Journal of General Psychology*, **129**, 3, 238–56.

Ware, J. E. Jr and Sherbourne, C. D. 1992. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Medical Care*, **30**, 6, 473–83.

Welte, J. W. and Russell, M. 1993. Influence of socially desirable responding in a study of stress and substance abuse. *Alcoholism: Clinical and Experimental Research*, **17**, 4, 758–61.

Wijndaele, K., Matton, L., Duvigneaud, N., Lefevre, J., Duquet, W., Thomis, M., De Bourdeaudhuij, I. and Philippaerts, R. 2007. Reliability, equivalence and respondent preference of computerized versus paper-and-pencil mental health questionnaires. *Computers in Human Behavior*, **23**, 4, 1958–70.

Women's Health Initiative Study Group 1998. Design of the Women's Health Initiative clinical trial and observational study. *Controlled Clinical Trials*, **19**, 1, 61–109.

Wrobel, A. J. and Shapiro, N. E. 1999. Conducting research with urban elders: issues of recruitment, data collection, and home visits. *Alzheimer Disease and Associated Disorders*, **13**, supplement 1, S34–8.

Address for correspondence:

Barton W. Palmer, Department of Psychiatry, University of California, San Diego, 9500 Gilman Drive 0603v, La Jolla, CA 92093-0603, USA.

E-mail: bpalmer@ucsd.edu