Differential Sensitivity to Administration Format of Measures of Attitudes Toward Older Adults

Edward Helmes, PhD,¹,² and Alistair Campbell, PhD²

Purpose: Reluctance to reveal sensitive or socially undesirable attitudes has posed a problem for measurement of personal attributes such as attitudes toward older people. These have long been documented to be negative and likely arise both from fears of one’s own aging and the modern societal emphasis on youth. In order to increase our knowledge about the measurement of attitudes toward older people, we compared the administration of attitude measures toward older people by computer and conventional paper-and-pencil methods. Design and Methods: We contrasted the responses of 60 university undergraduates (mean age 24.3 years, SD 8.51; 68% female) to five traditional paper-and-pencil format attitude questionnaires toward older people with responses to the same questionnaires made using an Internet delivery and response mode. Results: Results showed that more negative attitudes were revealed using the computer-based Internet response format for all scales. Only two scales did not show significant differences between formats. Implications: Future research on attitudes toward older people should be aware that results are dependent not only on the particular scale that is used but also on the format of administration, with more negative attitudes revealed with computer administration.

Key Words: Stereotype, Ageism, Older workers, Internet-based assessment

Recent years have seen self-report using paper-and-pencil questionnaires become the dominant form of attitude and personality assessment (Kaplan & Sacuzzo, 2006) for virtually all ages of participants. One issue that limits the utility of self-report data for some purposes is the extent to which the material being assessed is personally sensitive or socially undesirable (Newman et al., 2002), such as negative attitudes toward groups in society. Most approaches to assessing attitudes toward older adults rely upon self-report methods (Kite & Wagner, 2002), and results of the use of such scales generally reflect negative attitudes.

Such negative attitudes are less likely to be disclosed unless circumstances are anonymous. Being predominantly negative, attitudes toward older adults are likely to be among the attributes that are best addressed anonymously. The review by Hess (2006) summarizes research on attitudes toward older adults, which are generally negative. He also concludes that implicit methods of assessing attitudes generally reveal more negative attitudes than conventional attitude measures, a point also made by Levy and Banaji (2002). Traditionally, some supervision of individuals while completing self-report, paper-and-pencil instruments has been the norm, resulting in a degree of social interaction between the respondent and the supervisor and also a lack of true anonymity of any responses.

When questionnaires are administered by computer, such as with many implicit methods of attitude measurement, it is likely that such human supervision is used less frequently or is even nonexistent (Tourangeau & Smith, 1996), thereby resulting in greater confidence of true anonymity among respondents.

Benefits of self-report assessment are its relative privacy, potential anonymity, and the elimination of variations due to interviewers (Epstein, Barker, & Kroutil, 2001). In addition, large numbers of participants are relatively easily obtained, leading
to significant economies in assessment costs. Such economies can be even more substantial for Internet-based assessment (Fraley, 2007). Furthermore, obtaining information nationally or internationally becomes economically viable for more researchers. Durant and Carey (2000) concluded that the use of the paper-and-pencil self-report format can lead to increased reports of sensitive behaviors, including such behaviors as alcohol use (Aquilino & Lo Sciuto, 1990; Hochstim, 1967) and illicit drug use (Aquilino, 1994; Schober, Caces, Pergamit, & Branden, 1992; Turner, Lessler, & De Vore, 1992).

As might be expected, the perceived privacy of the assessment circumstances is highly relevant to one’s willingness to disclose negative attitudes. For example, family members may perceive a lack of privacy if other members of the family have the opportunity to see their responses (Epstein et al., 2001). Bradburn and Sudman (1979) concluded that, for honest self-report of sensitive information, most respondents needed evidence that there was no possible link between their identity and their responses. Another important element is the credibility of the purpose and person conducting the assessment (Weinhardt, Forsyth, Carey, Jaworski, & Durant, 1998). If the information is to be used in a legitimate prosocial manner, people are more willing to disclose sensitive information, regardless of the administration format of the questions.

At the same time, the traditional paper-and-pencil format does have disadvantages, including the inability to clarify questions (Tourney & Smith, 1996), a serious issue in samples in which there is a significant number of respondents with limitations in literacy. With the traditional format, respondents may read through the entire questionnaire before responding to the first item and may also omit items, which may be undetected until it is too late to be corrected, such as with questionnaires that are returned by post.

Wise and Kingsbury (2000) have pointed out that designers of computerized testing software need to consider issues such as the option of reviewing previous responses and that increased flexibility of options will optimize the quality of the assessment (Fraley, 2007). Simply providing or not providing the option of reviewing previous responses is a matter that needs consideration depending on the purpose of the assessment.

With advances in computer technology, computerized applications of conventional assessment instruments have become more common, despite concerns being raised about the equivalence of the two formats (Naglieri et al., 2004; Schmidt, 1997). The word processing of the basic paper-and-pencil forms, optical scanning of self-administered forms, computer-assisted telephone interviewing, and computer-assisted test administration have all become more widespread in use. The administration of questionnaires by computer has several distinct advantages. The elimination of several forms of administration errors, the immediate transfer of raw data into statistical packages for analysis, and better standardization of administration procedures have all contributed to its popularity.

Suler (2001) has pointed out that people are much more likely to disclose sensitive information in situations where that information is being collected by a computer without human supervision. In an early meta-analysis of some 39 studies comparing computer administration to pencil-and-paper administration of 100 various measures, Weisband and Kiesler (1996) found that computer-based administration consistently led to increased self-disclosure. In a more recent study, Joinson, Woodley, and Reips (2007) found that nondisclosure to an Internet-based questionnaire was related to how subjects accessed the survey and whether information was considered personally sensitive. There is also a growing body of evidence that Internet-based communication frequently involves more self-disclosure of both positive and negative views, values, and experiences than conventional paper-based methods (Barak & Gluck-Ofri, 2007; Lai-yee Ma & Leung, 2006). These results suggest that Internet-based measurement of attitudes toward older adults may be more negative than with conventional paper-and-pencil administration.

In assessing attitudes toward older people, the dominant method of assessment has been paper questionnaires (Kite & Wagner, 2002). To date, we are not aware of any studies that have explicitly contrasted traditional and Internet-based assessment of attitudes toward older adults. This issue is becoming increasingly important with expanding interest in assessing attitudes toward older adults (Hess, 2006; Palmore, 1990). Various psychological factors, such as authoritarianism, rationalization, and lack of information, have all been identified as contributing to the reported negative attitudes toward older people (Palmore). The recent compilation by Nelson (2002) summarizes the current state of much research on such ageist
attitudes, including a summary of attitudes toward older workers (McCann & Giles, 2002).

The primary purpose of this study was to contrast scores on instruments designed to assess attitudes toward older adults obtained using the traditional paper-and-pencil format compared with a computer-based administration of the same questionnaires. Scales were selected to assess both older adults in general and older workers. We hypothesized that the provision of greater anonymity through computer-based administration would lead to more negative attitudes being revealed than with the paper-and-pencil administration.

Methods

Participants

Participants were recruited from the undergraduate population of a Queensland (Australia) University. Initially, 90 students volunteered, but 30 were excluded due to excessive missing data (including failure to complete both assessments, which was the most common cause), leaving a total of 60 who completed all phases of the study. Of the total, 39 were recruited through a departmental subject pool, which provided course credit in exchange for participation. The remaining participants were volunteers recruited through a poster campaign and contacts made during lectures.

The majority were female (68%) and between the ages of 17 and 25 years (72%), with a mean age of 24.3 years (SD 8.51). The great majority (81%) were Caucasian, with 9% of Aboriginal and Torres Strait Island background, and the remainder being “other.”

Materials

Five paper-and-pencil self-administered questionnaires were converted to computer-based format using the Survey Crafter Professional (Survey Crafter, Inc., 2005) software package. The five scales were administered in the same order in both formats, as given below.

The Beliefs about Older Workers Questionnaire (Hassell & Perrewe, 1995) is a revised version of the 1952 “Attitudes Toward the Employment of Older People” Scale of Kirchner, Lindbom, and Paterson (1952). Low scores reflect more negative attitudes toward older workers. Rosen and Jerdee (1976) developed a set of adjectives relating to the potential of older people in general for development, performance capacity, and stability and interpersonal skills. Their scale uses a 10-point scale ranging from 0 (not at all accurate) to 9 (very accurate) to rate the extent to which the adjectives describe older people with lower scores reflecting more negative attitudes. Sample adjectives included versatile, accurate, accident prone, emotional, and sarcastic. Kogan (1961) described a scale using a 7-point Likert scale to assess both negative and positive attitudes toward older people. Odd-numbered items are scored for negative attitudes with high scores corresponding to more negative attitudes. Even-numbered items are associated with positive attitudes, and negative attitudes are indicated by low scores. The scale developed by Tuckman and Lorge (1952) to assess attitudes toward older workers uses a yes-no format with greater endorsements indicating more negative attitudes. The final scale used was the Attitude Toward the Employment of Older People Scale of Kirchner and Dunette (1954) that uses a 5-point Likert scale, revised from the earlier version that was also amended by Hassell and Perrewe. Low scores on this scale indicate more negative attitudes.

Procedure

The James Cook University Human Research Ethics Committee reviewed and approved the study prior to data collection. Respondents were randomly allocated into two groups, with order of administration (paper-and-pencil first or second) counterbalanced between the two groups. The administration instructions only indicated that the purpose was to investigate ways of thinking about older adults and older workers. The paper-and-pencil administration was conducted in a classroom in a single group, while the computer-based administration was conducted through a specialized contact system in which students were given a Web address and asked to log in individually at their convenience. The group that completed the Web version first was given the access codes and instructions for logging in by e-mail about a week prior to the paper-and-pencil administration, with access terminating a week prior to the scheduled session for the paper-and-pencil administration. Similarly, the group completing the Web administration second received access information a week following the traditional administration and could not access the Web version prior to that date.
Results

Table 1 reports means, standard deviations, and coefficients alpha for the six measures (both negative and positive for the Kogan scale). All values of coefficient alpha were in excess of .8 for both administration formats, with the exception of the Kirchner scale in both cases.

Within-subject t tests were conducted for each measure. Table 1 also reports values of the t tests, 95% confidence intervals for the difference in scores between administration formats, and effect sizes for those differences. Effect sizes were calculated using the pooled variance estimate from each occasion as being more conservative than using the variance of the difference between occasions (Dunlop, Cortina, Vaslow, & Burke, 1996). There were no differences between administration formats for the Hassell and Perrewe scale (p = .59) and the Kogan positive scale (p = .10). Of the remaining three scales for older workers, more negative attitudes were expressed for the computer-administered Tuckman–Lorge scale (p < .001), the Rosen–Jerdee scale (p < .001), and the Kirchner scale (p = .01). For a measure of attitudes toward older adults in general, the Kogan negative scale also showed significant differences between occasions (p = .02). Note that negative attitudes are expressed in opposite directions with the Tuckman–Lorge and Kogan negative scales to the Rosen–Jerdee and Kirchner scales. Findings were consistent in that in all cases of significant differences between administration formats, more negative attitudes were expressed using the computer-administered format. Indeed, the direction of the differences was the same for the two scales that did not show significant differences between administration formats.

Discussion

The conditions of the two assessment occasions contrasted the common differences between paper-and-pencil and computer-based assessment. The paper-and-pencil assessment was conducted in a classroom setting, with individuals potentially able to observe others’ responses, and with clear potential for social interactions both with other respondents and with the administrator. In contrast, individuals completed the computer-based assessment at their convenience with as much privacy and anonymity as they desired. The effect sizes for the differences between formats ranged from just less than one half a standard deviation to almost 2 SDs for the Tuckman–Lorge scale. These are medium to large effect sizes according to the commonly used guidelines to labels for effect sizes provided by Cohen (1988). Clearly, the scales differed in their sensitivity to the format of administration, suggesting that researchers need to consider carefully their selection of measures and administration formats.

Internal consistencies were equivalent across the forms, with five of the six measures showing good reliabilities in this sample. The only exception was the Kirchner scale, which had internal consistency reliability coefficients below .7 for both forms of administration. This result suggests that this scale is less suitable for measuring attitudes toward older workers than the other measures. The Kirchner scale was also the last one administered, and it is possible that a form of fatigue had set in by the time participants reached it, having completed four other scales by then with very similar content. Randomizing or counterbalancing the order of administration of the scales would have reduced the

Table 1. Means, Standard Deviations and Alpha Reliabilities for Attitude Measures (n = 60)

<table>
<thead>
<tr>
<th>Measure</th>
<th>Number of items</th>
<th>Paper administration</th>
<th>Computer administration</th>
<th>Value of t test</th>
<th>95% CI</th>
<th>Effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>M</td>
<td>SD</td>
<td>Alpha</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Hassell and Perrewe</td>
<td>36</td>
<td>108.2</td>
<td>13.30</td>
<td>.901</td>
<td>107.1</td>
<td>16.35</td>
</tr>
<tr>
<td>Rosen–Jerdee</td>
<td>65</td>
<td>350.8</td>
<td>45.19</td>
<td>.920</td>
<td>314.1</td>
<td>49.11</td>
</tr>
<tr>
<td>Kogan negative</td>
<td>17</td>
<td>48.9</td>
<td>10.03</td>
<td>.827</td>
<td>51.7</td>
<td>12.54</td>
</tr>
<tr>
<td>Kogan positive</td>
<td>17</td>
<td>71.9</td>
<td>9.53</td>
<td>.813</td>
<td>69.2</td>
<td>11.16</td>
</tr>
<tr>
<td>Tuckman–Lorge</td>
<td>51</td>
<td>16.3</td>
<td>10.73</td>
<td>.943</td>
<td>36.5</td>
<td>10.42</td>
</tr>
<tr>
<td>Kirchner</td>
<td>24</td>
<td>71.1</td>
<td>6.26</td>
<td>.611</td>
<td>67.3</td>
<td>5.99</td>
</tr>
</tbody>
</table>

Notes: CI = confidence interval. All within-subject t tests have 59 df.
*p < .05. ***p < .0001.
likelihood of such a possible fatigue effect influencing only one scale.

There was complete consistency across the five scales and six measures used. In four of the six cases, significantly more negative attitudes toward older people were expressed in the computer-based form than in the paper-and-pencil form, and the pattern was the same for the two scales that did not show statistically significant differences. Comparison of scores obtained in this study with previous work shows that scores on the two Kogan (1961) scales were more negative than those in the report by Stewart, Giles, Paterson, and Butler (2005) for a group of New Zealand health care students, particularly for the computer-administered version. This was also the case for the Tuckman and Lorge (1952) scale, but here, the value obtained for the mean of the paper-and-pencil administration was in the range that was originally reported. In contrast, the means for the Hassell and Perrewe (1995) scale under both conditions were notably more positive than those originally reported.

Previous studies have found that people seem to respond to computers with more disclosure of their “real” feelings when they feel that the administration is both anonymous and does not involve sensitive personal information (Barak & Gluck-Ofri, 2007; Joinson et al., 2007). The higher negative scoring for the Internet-based questionnaires may therefore reflect a more accurate picture of people’s perceptions. This suggests that the anonymity associated with computer-based administration of tests may lead to more honest answers or, at least, answers less affected by the social desirability that might be aroused with conventional administration with accompanying human interaction. Although Chuah, Drasgow, and Roberts (2006) found that Internet and traditional administrations were equivalent using a variety of criteria, they also reported that there was a significant difference between the correlations of emotional stability and risk of substance use between Internet format and traditional versions. This was also the case for the Hassell and Perrewe (1995) scale under both conditions.

One suggestive finding of this study was the generally more negative attitudes expressed on the scales for older workers than those on the scales for older adults in general. The effect sizes for the difference between the paper-and-pencil and the Internet-based scores for the Tuckman–Lorge and Rosen–Jerdee scales were particularly large, and with the exception of the Hassell–Perrewe scale, the effect sizes for the scales about older workers were all larger than those for the Kogan scales. Further research might explore this issue in more depth using different measures and experimental designs than those in this study to determine if older workers are indeed perceived more critically than older adults in general. The within-subject design used here is more powerful in detecting differences between administration formats than a design in which participants are randomly allocated to two different groups, one completing paper-and-pencil measures and one receiving questionnaires by computer-based administration. The latter design does avoid the problem of individuals remembering their responses and duplicating them in order to maintain consistency during the second administration. At the same time, with a week between administrations and over 200 items, it is unlikely that participants remembered their responses, but they may have recalled their overall attitude and attempted to maintain their consistency. This effect would, however, act to reduce any differences between the two administration formats and lower the likelihood of significant differences.

Within the current design, there is no way of knowing whether the negativity that was observed was indeed a reflection of a more “honest” view, but this more likely than not given the findings of generally negative attitude toward older people in society (Hess, 2006; Kite & Wagner, 2002). It does raise an interesting question, however, as to whether computer-based administration of questionnaires (whether that be with the Internet or local individual computers) leads to a reduction in more general social desirability effects. Chuah and colleagues (2006) used adjectives that likely were comparatively neutral in desirability and did not use highly undesirable or sensitive content. Use of questionnaires about aging may be particularly relevant to this line of inquiry because of the potentially large impact of social desirability on the expression of negative values in this area.

Another implication of our results is that researchers should exercise caution when comparing studies that have used computer-based administration of these measures to those that have used paper-and-pencil administration. Our findings suggest that the results may not be strictly comparable between the two administration formats.
There would thus need to be some explicit check as to whether the psychometrics of a particular instrument were affected by computer-based administration. This is a more general caution that can be applied to the comparison of computer-based and paper-and-pencil versions of many attitude measures. Our study suggests that the validity of computer-based administration, through comparison with equivalent paper-and-pencil forms, needs to be established before they can be used as straightforward substitutes.

Acknowledgments

We thank J. Bendzulla, L. de Jager, L. Heilbronn, and A. Weightman for their efforts in administering the measures and collecting the original data.

References


Received January 2, 2009
Accepted March 30, 2009
Decision Editor: William J. McAuley, PhD

Vol. 50, No. 1, 2010 65