

Faculty of Commerce

Faculty of Commerce - Papers

University of Wollongong

Year 2007

How Demographic Characteristics Affect
Mode Preference in a Postal/Web
Mixed-Mode Survey of Australian
Researchers

K. Diment*

S. Garrett-Jones†

*University of Wollongong, kdiment@uow.edu.au

†University of Wollongong, sgarrett@uow.edu.au

This is an author version of an article that will be published by Sage as: Diment, K & Garrett-Jones, S, How Demographic Characteristics Affect Mode Preference in a Postal/Web Mixed-Mode Survey of Australian Researchers, Social Science Computer Review, Fall 2007, 25(3). Journal information available here.

This paper is posted at Research Online.

<http://ro.uow.edu.au/commpapers/304>

How demographic characteristics affect mode preference in a postal/web mixed-mode survey of Australian researchers

Kieren Diment¹ and Sam Garrett-Jones¹

¹Centre for Leadership and Knowledge Management, School of Management and Marketing, University of Wollongong, NSW 2522, Australia. kdiment@uow.edu.au

Abstract

Early promise for the internet as a tool to make social research questionnaires easier and cheaper to deliver is not fully realized. This study reports a mixed-mode survey of 1100 Australian researchers. When respondents were given the choice to complete either web or paper-based versions of a questionnaire the majority chose the paper-based mode. Web respondents were more likely to be young, male, middle ranking and working in information technology related sectors. We highlight the need to determine how far alternate delivery modes increase response rates. For mixed-mode surveys to be financially and methodologically worthwhile we propose that the initial sample size be at least 1000 individuals, this figure depending on the demographic characteristics of the sample.

Keywords: web surveys; survey methodology; internet; demographic differences; workplace research

Introduction

Use of the internet as a questionnaire delivery mechanism appears an attractive option. As respondents effectively act as their own data-entry clerks, there is a potential for cost savings and. Computer-based validation of user input can increase data quality. However, respondents may not be enthusiastic about responding to questionnaires over the internet, and so mixed-mode questionnaires where respondents have the choice of a paper, web, fax or possibly telephone based questionnaire have received attention (Carini et. al. 2003, Leece et. al. 2004). The concern is that the mixed-mode approach might bias the survey sample.

Investigating the management of collaborative research we conducted a survey of employees from higher education, government research and other public sector organizations in Australia. All were participants in the Cooperative Research Centre (CRC) program of cross-sector collaborative research and development organizations (DEST, 2004).

The penetration of the internet within the Centres is close to complete. The survey population was highly educated and presumably relatively homogeneous in their computer literacy. This is important, since research has shown that, in other applications such as web-based personnel recruitment (Hinton 2003), the propensity for internet use depends on access to and capacity to use the technology. Our sample provides a good milieu for assessing the kinds of individuals who will choose web based questionnaire completion over a postal survey.

It is clear from the literature that , differences in web and paper based delivery mode do *not* usually alter responses enough to affect research hypotheses (Carini et al., 2003; Layne et al., 1999). Client side or server side scripting can be used to validate the responses to ensure that the data quality is high (Couper, 2000). However,

it is apparent that web based questionnaires are less popular than postal or fax surveys (Leece et al., 2004; Schonlau et al., 2003). In order to maximize returns a mixed mode strategy of paper and web delivery is potentially attractive. Additional costs of a mixed-mode approach include debugging, ensuring integrity of data storage, provision of access control to prevent abuse, and server infrastructure.

Although the mixed mode approach can realize significant savings, Schonlau et al., (2003) indicate that sample sizes need to be large. How large must be determined by sample demographics. However, our understanding of these demographic factors and their influence is limited. To better define the cost/benefit analysis used to determine if a mixed mode strategy is appropriate, we examine several demographic variables to determine which are associated with a preference for web based completion.

Some demographic aspects are evident. While the gender gap in internet usage is narrowing, females remain less enthusiastic and intense users of the internet (Ono and Zavodny, 2003). A sample biased towards females is *less* likely to benefit from the mixed-mode approach than a gender balanced sample. The great majority (82%) of our respondents were male.

Although all our respondents were employed in the public sector, the Centres are categorized according to broad industry-related sectors: Medical, Mining, Agriculture, Information and Communication Technology, Manufacturing and Environment (DEST 2004). National data show that the sector containing scientific research organizations (Property and Business Services) has the highest internet penetration (ABS 2005). One of our aims is to identify whether individuals within particular industry-related research sectors are more likely to prefer a web based questionnaire over a parallel postal version.

Method

We sought a broad sample of individuals from the public sector with a research interest in the CRC program. We could not identify a census of relevant workers as many work part time on CRC projects while employed by their 'parent' organization. The DEST releases statistics on the number of full time equivalent staff for each CRC, but not information on the number of individuals. This is due to the distributed multi-partner, multi-site nature of the CRCs, and the mobility of individuals within and between CRCs and partner organizations. Because of this, we describe CRCs as 'cryptic' organizations.

Instead we identified a 'convenience' sample of these individuals. Many CRCs release their annual reports as Portable Document Format (Adobe Acrobat) documents on their websites. We obtained every current annual report available in this format (28 of 73 total CRCs) and converted them into plain text for further processing. A simple computer program (using the Perl programming language and a third party libraries) was used to identify the names of individuals and institutions in the text. Around 2000 names and institutional affiliations were found of which 1100 were associated with postal addresses found with the 'Google' search engine.

A nine page questionnaire was sent out to all 1100 individuals in the sample. A covering letter invited them either to fill out the questionnaire online or to complete the booklet. The security of the online questionnaire was protected by a password, and the (same) user name and password was provided to all potential respondents in the covering letter. This prevented casual abuse of the questionnaire page but would not have prevented more serious hacker attempts. The IP (internet) address of the computer that completed the questionnaire; the URI (web address) of the page that requested that data be inserted into the survey database (the referrer), a time stamp

showing when the data was inserted into the survey database were collected as well as the survey questions to help verify the integrity of the data collected over the web. This made it possible to identify multiple completions by the same respondent, for example. The server ran Red Hat Linux; the database store was a Mysql server on the same computer.

To ensure that the online survey and the printed version were as similar as possible, the structure and content of the questionnaire was described in an Extensible Markup Language (XML) document. This XML document was transformed into a LaTeX source file. LaTeX is a scientific typesetting system that will easily produce press ready output and is easy to produce programmatically. All software used to produce the survey materials was free open source. The HTML code for the web based survey was also produced programmatically from the XML file. The HTML was written to be viewable in any web browser, including text only browsers. To ensure cross-platform delivery only server side validation of the user input was performed. The CGI script for validating the input from the survey and inserting the data into a database was hand written in Perl. There was one disadvantage to the mixed-mode approach. While the CGI script did not permit incomplete questionnaires to be entered into the survey database, responses from the paper questionnaires necessarily allowed for questions to be left blank. With an anonymous questionnaire there was nothing that could be done to rectify this.

Questionnaire Structure

The survey questionnaire comprised six parts: personal details, employer and CRC details; a series of propositions about the respondent's experience with the CRC program requiring on a six-point forced choice Likert scale (Strongly Disagree,

Disagree, Slightly Disagree, Slightly Agree, Agree and Strongly Agree). A 'Not Applicable' response was also included. Each section had an open ended question at the end.

Data Analysis

The respondents were split into two groups: web respondents and paper respondents. We compared the two groups by gender, highest qualification, year this qualification was obtained (a proxy for age and seniority), employer institution (primarily government or university), job type and length of involvement with the CRC. We also investigated the influence of the industry-related sector on response mode. Chi-square (χ^2) goodness of fit tests were used to determine if the percentage of web respondents differed significantly across the groups compared to postal respondents.

Results

Of the 1100 questionnaires distributed, 77 (or 6.4%) were returned as 'unknown at this address' leaving 1023 potential valid respondents. Of these, 370 responses were received, a response rate of 36% of the total valid sample. A total of 295 respondents completed the questionnaire using the enclosed paper survey form (80% of respondents) while 75 (20%) responded over the web. Even this sample of 'internet savvy' researchers favored the traditional 'paper-based' method by a ratio of four to one.

We then examined how web respondents and paper respondents differed in their demographic characteristics. We present the significant demographic variables (i.e. those which discriminate between the two groups of respondents). Table 1 shows the proportion of respondents by response mode for the significant demographic

variables: gender, job type, industry-related sector and year of highest qualification. No significant difference was found in mode preference for highest academic qualification ($\chi^2_{df=3} = 4.78$ $p > 0.05$), employer ($\chi^2_{df=5} = 1.97$ $p > 0.05$), and length of involvement with the CRC ($\chi^2_{df=3} = 1.12$ $p > 0.05$).

TABLE 1 ABOUT HERE

Table 1a shows the proportion of responses to the survey by response mode by gender. Males were more likely to complete the questionnaire over the web ($\chi^2_{df=1} = 4.25$ $p < 0.05$). Proportionally, more than twice as many males (22%) chose to complete the web-based questionnaire than did females (11%).

Table 1b shows the effect of respondent's job type on response mode. There is a significant difference between the groups ($\chi^2_{df=7} = 14.66$ $p < 0.05$). Individuals in middle ranking positions - Other research, Technical officer and Middle management - were more likely to choose web based completion of the survey than were the other groups. Even so, in none of these groups did a majority chose web-based completion over paper-based. This is surprising as online reporting is common in universities and research funding agencies in Australia.

Table 1c shows the industry-related sector of the CRC by response mode. Note that industry sector was unavailable for 8 respondents. While there is no significant difference across all the groups ($\chi^2_{df=5} = 9.73$ $p = 0.08$) the Information Technology and Manufacturing Technology sectors show a greater tendency to respond via the web. However, even in these groups, more than half of the respondents chose to complete the postal version of the questionnaire.

Information on year of highest qualification provides a broad proxy for age and seniority. As Table 1d shows for the 329 responses with this data available, more senior staff preferred the paper based questionnaire ($\chi^2_{df=4} = 9.7, p < 0.05$) and the younger the researcher the greater the likelihood of choosing web completion.

Discussion

The rationale for employing a mixed mode survey includes achieving a greater reach for the survey (increasing the response rate), lower cost (through greater degree of automation) and achieving greater accuracy and completeness in data capture. The need to ensure high quality samples of potential respondents and the growth of ‘spam’ means that even where email contact is feasible, post, fax or telephone should remain the preferred method for inviting responses. However, the cost savings for data-entry, and improved data quality, means that a web based option is potentially attractive.

Response Rate

Our understanding of mode preference in mixed mode surveys is surprisingly limited. Couper (2000) cites several studies, one of which found that only 8% of respondents chose the web in a mixed-mode survey. This lack of popularity of is borne out in a randomized study of physicians asked to complete a survey either via the web or by post (Leece et al., 2004). The postal method was significantly more popular although both response rates were high (45% for the web respondents and 58% for postal). Our survey, with a 20% completion over the web, accords with the results obtained by Schonlau et al. (2003) who found 23% of respondents elected to use the web..

These studies perhaps indicate that with the increase in use and availability of internet connections at work and at home, the web is becoming a more popular option.

Our finding that the more recently qualified researchers favored the web suggests that future response rates to web based surveys will continue to increase. Unlike Schonlau et al., we sent the postal survey initially to all potential respondents with the invitation to complete the questionnaire by either mode. Schonlau et al. provided subjects with an invitation to complete their questionnaire over the web, with the option of requesting a hard copy of the questionnaire (although subsequent follow-ups included a hard copy). This indicates that making the postal survey less accessible is unlikely to have a positive effect on the total response rate or the web based response rate. We cannot say from our study whether the mixed-mode survey improved the response rate over a paper based survey alone. However, giving potential respondents an alternative way to complete questionnaires may increase the overall response rate, as some individuals who would not have completed a postal version of the questionnaire may elect to complete an electronic version. This question deserves further research.

Bias in mixed mode surveys

Does the option of a web-based response introduce bias into the final data? Our results show that four factors had a significant effect on mode preference.

(a) *Gender*. In our sample, Internet response is more than twice as attractive to male respondents than to females (c.f. Ono and Zavodny, 2003).

(b) *Youth*. Taking 'year of highest qualification' as a proxy for age we see an inverse correlation between preference for a web-based response and the age of the respondent (Table 1d). Fewer than 15% of those who qualified in the 1970s chose the web-based response, compared to nearly 33% of those who qualified since 2000. This result is not related to length of involvement with the CRC program, which had

no significant effect on mode. We are dealing with skilled population that has very high level of internet use and have therefore eliminated confounding factors such as differential access to technology and educational attainment.

(c) *Job type*. Senior researchers and senior managers were less likely to elect for web based completion than all other job types apart from postgraduate students who for some reason were also reluctant web-users (Table 1b),

(d) *Industry-related sector*. While it is not surprising that the Information Technology CRC workers were the most enthusiastic users of the web based questionnaire (Table 1c), we are unsure why Manufacturing-related CRC employees were more attracted to the web survey than were Medical or Agriculture researchers, for example. Further work is needed to explain why other respondents in this 'internet intensive' scientific research sample were more reluctant to use the web.

Cost effectiveness

The mixed-mode approach with a web option is only likely to realize cost savings where the sample size is relatively large. Exactly how large is an open question as the answer is mediated by demographic and methodological factors. If the typical response rate to the web based components of mixed mode surveys is low (at 10-20% of respondents) and assuming an overall survey response rate of approximately 30% for both modes, mixed mode surveys probably make economic sense only in situations where the number of initial subjects approaches 1000 or higher. For smaller samples, where intensive follow-up is possible, a single mode of delivery is probably the better approach. For short questionnaires, where data entry costs are low, a mixed mode survey is efficient only for very large samples. Our survey with 75 web respondents was at the margin of being cost effective as a mixed-

mode survey. Using the same methods in this study for subsequent surveys would result in substantial cost savings due to the amortization of development costs.

The minimum sample size for an effective mixed mode survey will decrease as the population of interest becomes more biased towards males, younger and/or less senior individuals, and research staff in information technology or possibly manufacturing related sectors. Our study suggests that age and gender may well be the more important demographic factors.

In conclusion, we show that even when surveying a highly educated, apparently homogeneous population of research scientists, few prefer to respond via the internet when given a choice. Further, there are significant differences between those respondents who choose the internet and those who prefer a paper based survey questionnaire which may bias the survey. However, this bias may be acceptable where use of the web enhances the scope and quality of the data received.

Acknowledgements.

This work was funded by an Australian Research Council Discovery Grant DP0211298, 'Managing the risks of cross-sector R&D collaboration'.

References

- ABS (2005) "Business Use of Information Technology" Australian Bureau of Statistics catalogue no. 8129.0 Commonwealth Government, Canberra, Australia
- Carini, R. M., Hayek, J. C., Kuh, G. D., Kennedy, J. M., and Ouimet, J. A. (2003). College student responses to web and paper surveys: does mode matter? *Research in Higher Education* **44**, 1-18.
- Couper, M. P. (2000). Web Surveys: A Review of Issues and Approaches. *Public Opinion Quarterly* **64**, 464-494.
- DEST. (2004). "CRC Directory." Commonwealth Department of Education Science and Training, Canberra, ACT, Australia.
- Layne, B., DeCristoforo, J. R., and McGinty, D. (1999). Electronic versus traditional student ratings of instruction. *Research in Higher Education* **52**, 1-26.
- Leece, P., Bhandari, M., Sprague, S., Swiontkowski, M., Schemitsch, E., Tornetta, P., Devereaux, P., and Gyuatt, G. (2004). Internet Versus Mailed Questionnaires: A Randomized Comparison (2). *Journal of Medical Internet Research* **6**, e30.
- Ono, H., and Zavodny, M. (2003). Gender and the Internet. *Social Science Quarterly* **84**, 111-121.
- Schonlau, M., Asch, B., and Can, D. (2003). Web Surveys as Part of a Mixed-Mode Strategy for Populations that Cannot Be Contacted by E-Mail. *Social Science Computer Review* **21**, 218-222.

Kieren Diment is Research Assistant for the Centre for Leadership and Knowledge Management at the School of Management and Marketing, University of Wollongong. His time is divided between providing research support for individual researchers, and providing web based collaboration tools for researchers. kdiment@uow.edu.au

Sam Garrett-Jones has a PhD in the natural sciences and is an Associate Professor in the School of Management and Marketing at the University of Wollongong, NSW, Australia. sgarrett@uow.edu.au

a. Gender

	Percentage (Count)			χ^2
	Paper	Web	Total	
Male	77.5 (235)	22.5 (68)	100 (303)	4.25(df=1)*
Female	89.1 (59)	10.9 (7)	100 (66)	
Total	79.5 (291)	20.5 (75)	100 (369)	

b. Job Type

	Percentage(Count)		χ^2
	Paper	Web	
Senior Manager	27 (79)	15 (11)	14.7(df=7)*
Senior Researcher	44 (129)	40 (30)	
Other Researcher	14 (40)	21 (16)	
Technical Officer	3 (8)	7 (5)	
Middle Management	2 (7)	7 (5)	
Postdoctoral	3 (10)	8 (6)	
Postgraduate	3 (10)	3 (2)	
Other	4 (11)	0 (0)	
Total	100(294)	100(75)	

c. Industry-related sector

	Percentage (Count)		χ^2
	Paper	Web	
Medical	11 (32)	4 (3)	9.73(df=5) ^o
Mining	16 (47)	12 (9)	
Agriculture	19 (56)	16 (12)	
IC&T	10 (29)	22 (16)	
Manufacturing	10 (28)	15 (11)	
Environment	33 (96)	30 (22)	
Total	100(288)	100(73)	

d. Year of highest qualification

	Percentage(count)		χ^2
	Paper	Web	
1950s and 60s	95(18)	5(1)	9.7(df=4)*
1970s	86(82)	14(12)	
1980s	75.5(74)	24.5(24)	
1990s	73(53)	27(20)	
2000s	67(37)	33(18)	
Total	72(254)	23(75)	

* $p < 0.05$ ^o $p < 0.1$

Table 1. Statistically significant demographic variables