Leading education and social research
Institute of Education
University of London

Data note

Response rates and potential response bias relating to the use of the paper self-completion questionnaire used as part of the 2008 sweep of the National Child Development Study

Jane Elliott and Matt Brown

2 December 2011

Centre for Longitudinal Studies
Following lives from birth and through the adult years
www.cls.ioe.ac.uk

CLS is an ESRC Resource Centre based at the Institute of Education, University of London

Data note on response rates and potential response bias relating to the use of the paper self-completion questionnaire used as part of the 2008 sweep of the National Child Development Study

Jane Elliott and Matt Brown

2 December 2011

The aim of this brief descriptive paper is to provide information on response to the paper self-completion questionnaire that formed part of the data collection for the eighth sweep of the 1958 British birth cohort study (also known as the National Child Development Study). This was carried out in 2008 when cohort members were aged 50. As reported by Brown (2008), the pilot for the self-completion questionnaire was very successful and resulted in a response rate of 92 per cent. This type of paper self-completion questionnaire is also a very cost effective way of collecting information from cohort members and is particularly suited to questions that are applicable to the whole sample. However, the pilot sample was not large enough to use for the analysis of which factors are associated with response to the paper self-completion questionnaire. This data note therefore uses information from the main survey of cohort members, at age 50, to examine which cohort member characteristics are associated with response to the self-completion questionnaire. This information is useful both for planning future similar modes of data collection, but also for understanding how differential response rates from different groups of cohort members may influence the results of analysis using variables from the self-completion questionnaire.

Those cohort members who were eligible to take part in the 2008 sweep of the 1958 cohort study were sent a paper self-completion questionnaire by post. They were asked to complete it in advance of their face to face interview. The questionnaire was A4 size and sixteen pages long, and was estimated to take approximately 20 minutes to complete. Questions focused on how cohort members spend their leisure time, how they rate their health, and whether their health impacts on their daily life and overall wellbeing. The questionnaire also included a fifty-item personality inventory (Goldberg, 1999). Where cohort members had not completed the questionnaire by the time the interviewer arrived, they were asked if they could spare the extra time at the end of the interview to complete it then and there. Where this was not possible, the interviewer arranged to return to pick up the completed questionnaire or left a free post envelope for the respondent to return it to the office themselves. In this way the response to the paper self-completion questionnaire was maximised. The fieldwork agency (NatCen) recorded that out of 9,790 productive interviews, 8,844 paper questionnaires were completed (90.3 per cent). There were no instances where a paper questionnaire was received but there was no main or partner interview. However, analysis of useable data from the self-completion questionnaire suggests a slightly lower response rate with 8,788 completed questionnaires returned (i.e. 89.8 per cent).

Modes of data collection and burden on respondents

It is well established in the survey methodology literature that there are four main steps involved when respondents answer survey questions:

- comprehension or understanding the question
- retrieval of requested information from memory
- evaluation of the salience of the retrieved information
- communication of the response (Tourangeau et al., 1984).

It is therefore likely that the channel or mode of questionnaire presentation will have an effect on the cognitive burden placed on respondents. In particular, in the case of paper self-completion questionnaires there is a demand for literacy. As Bowling (2005) has argued, probably the least burdensome method of data collection for respondents is the face-to-face interview because the respondent is only required to have basic verbal and listening skills and no reading skills are required. In addition, a well-trained interviewer can increase response and item-response rates, maintain motivation with longer questionnaires, and help clarify ambiguous questions.

In the 2008 sweep of the 1958 birth cohort study the fact that the main mode of data collection was a face to face interview is likely to have substantially boosted the response rate to the paper self-completion questionnaire. In particular, by sending the questionnaire before the main interview it was possible for the interviewer to collect it at the time of the interview, and to encourage the cohort member to complete the questionnaire if they had not already done so.

Response to the paper self-completion questionnaire by gender, partnership status and qualifications

It would be expected that those completing the questionnaires were not simply a random subgroup of those who participated in the main data collection, i.e. the face-to-face interview. We therefore investigate the characteristics of those who returned the questionnaire, compared with those who did not. As can be seen from Table 1, simple bivariate analysis suggests that cohort members who are male, not living with a partner and with no qualifications[i] are less likely to complete the questionnaire than other groups of cohort members. Using the Chi-squared test, all these bivariate associations were found to be significant at the $p<0.05$ level. However, it should be noted that differences between groups in their propensity to return the questionnaire are relatively small, such that even among those with no qualifications there were still high response rates with over 83 per cent completing the questionnaire.

Table 1: Response rates by characteristics of cohort members

Characteristic	Proportion completing paper questionnaire	Total N
Male	0.877	4822
Female	0.918	4968
No spouse or partner	0.874	2026
Lives with spouse or partner	0.906	7733
No qualifications	0.839	1096
Quals at NVQ level 1	0.895	1087
Quals at NVQ level 2	0.893	2491
Quals at NVQ level 3	0.903	1686
Quals at NVQ level 4	0.918	2998
Quals at NVQ level 5 or 6	0.929	425

Response to the paper self-completion questionnaire by cognitive capability

Given that there is a clear association between qualifications and response to the paper self-completion questionnaire it is helpful to look in more detail at the association between cohort members' measured cognitive capability at age 50 and the likelihood of them completing the paper questionnaire.

There were four measures of cognitive capability included in the age 50 sweep of the NCDS:

1. Word list recall (10 words)
2. Animal naming (number of animals named in a minute, range = 1-65)
3. Letter cancellation (number of Ps and Ws correctly marked, 0-65)
4. Delayed word list recall (10 words)

Further details of these assessments are provided by Brown and Dodgeon (2010).

A series of t-tests were carried out to investigate whether there are differences in the cognitive capabilities of those who completed the paper questionnaire in comparison with those who did not. As can be seen in Table 2 below, there are significant differences on each test such that those who did *not* return a self-completion questionnaire have lower cognitive capability scores than those who returned the questionnaire, but these differences are not large.

Table 2: Association of cognitive capability and return of self-completion questionnaire

Cognitive test	Self-completion paper questionnaire	N	Mean	Std. Dev	Std. Error Mean	Mean diff.	t	Sig.
Animal naming (Number of animals named)	Not completed	943	21.06	6.569	0.214	-1.355	-6.283	<0.001
	Completed	8706	22.41	6.258	0.067			
Letter cancellation (Number of Ps and Ws correctly marked)	Not completed	874	20.72	5.840	0.198	-0.986	-4.721	<0.001
	Completed	8568	21.71	5.889	0.064			
Word-list recall (Number of words recalled)	Not completed	943	4.95	1.984	0.065	-5.09	-8.071	<0.001
	Completed	8649	5.46	1.821	0.020			
Delayed word-list recall (Number of words recalled)	Not completed	943	6.22	1.604	0.052	-3.58	-7.038	<0.001
	Completed	8706	6.58	1.470	0.016			

The strongest association in Table 2 is between delayed word recall and whether the questionnaire was returned, with an associated t-value of -8.1. Figure 1 below shows the distributions of the delayed word recall scores for the two groups. It can be seen that although the shape of the distribution is slightly different for those who did not return the paper questionnaire – with fewer at the top end of the distribution and more at the lower end – the difference is not striking enough to make a major impact on models including only those with data on the paper questionnaire.

Figure 1: Distribution of scores on delayed recall task for those who returned or did not return the self-completion questionnaire

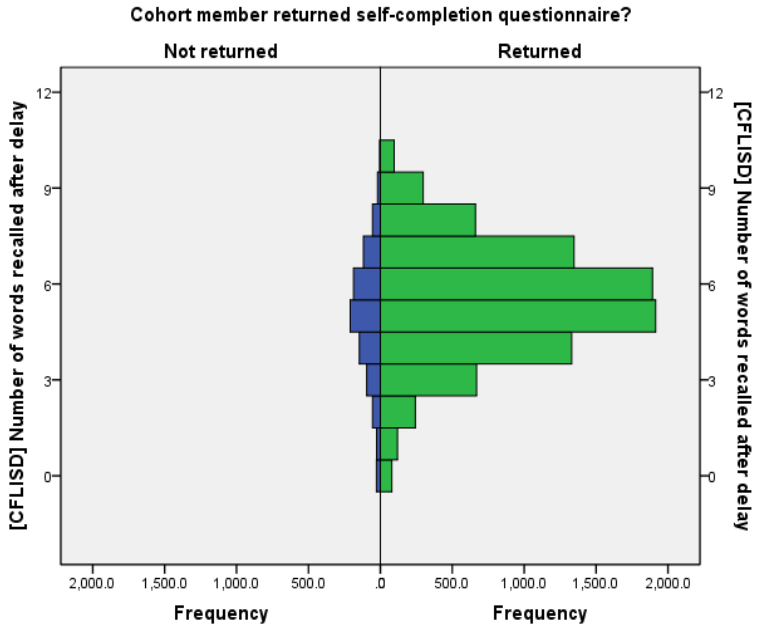
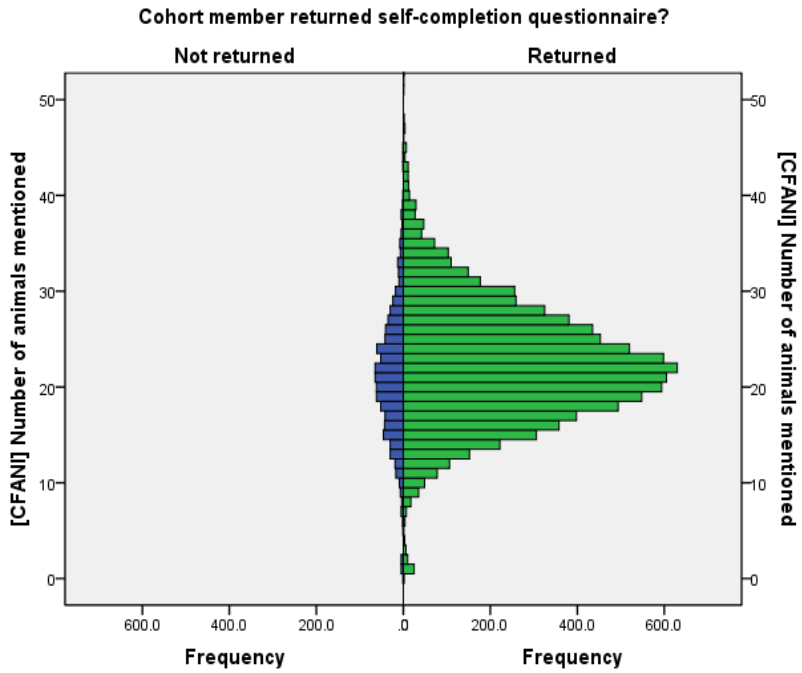


Figure 2 shows the animal naming score for the two groups. This also shows a different shaped distribution for those who did not complete and return the paper questionnaire, with a deficit of individuals with very high scores.

Figure 2: Distribution of scores on the animal-naming task for those who returned, or did not return the self-completion questionnaire



Response to the paper self-completion questionnaire by self-reported health status

Given that many of the questions included in the paper self-completion questionnaire focus on health, capability and wellbeing, it is important to discover whether the health status of those who completed the questionnaire is substantially different from those who did not complete the questionnaire. More specifically we also investigate whether those with impaired vision are less likely to return the questionnaire. As can be seen in Table 3 below, while there are associations between cohort members' self-reported health status and their propensity to return the paper questionnaire, the effects are very small such that the health of those returning the questionnaire is very similar to that of those who do not return the questionnaire. Using the Chi-squared test, associations were found to be significant between response and the variables in Table 3, with the exception of visual impairment, where numbers are much smaller. Most notable is the association between self-assessed health and propensity to complete the paper questionnaire with 92 per cent of those rating their health as excellent completing the questionnaire compared with 84 per cent of those rating their health as poor.

Table 3: Self-reported health and response to paper questionnaire

		Proportion completing paper questionnaire	Total N
Whether health limits daily activities	Yes	0.875	1532
	No	0.909	7145
Malaise score grouped	Low (0-3)	0.906	8216
	High (4+)	0.884	1411
Self-assessment of health	Excellent	0.920	1890
	Very good	0.910	3207
	Good	0.900	2841
	Fair	0.873	1239
	Poor	0.835	557
Visual impairment[ii]	Yes	0.877	57
	No	0.898	9733

Item non-response on the paper questionnaire and in the face to face interview

The level of non-response at item level is clearly another important indicator of data-quality (De Leeuw et al., 2008) and item non-response is generally reported as being higher in postal surveys than in face to face (and telephone) surveys (Harris et al., 1997, Brazier et al., 1992, Nicholaas et al., 2000).

The self-completion questionnaire contained 162 closed questions. Table 4 below shows the levels of missing data arising both from item non-response and errors (multiple responses to single response items) to the 162 closed questions.

Table 4: Missing responses / errors

Number of missing responses / errors	Proportion	Total N
None	0.785	6902
1 (0.6% of questions)	0.110	969
2 (1.2% of questions)	0.036	320
3 (1.8% of questions)	0.016	140
4 (2.4% of questions)	0.009	78
5-10 (3%-6% of questions)	0.019	166
11 or more (more than 6% of questions)	0.024	213

Four out of five respondents (79 per cent) completed all 162 items (with no errors) and a further one in ten (11 per cent) missed or made an error on just one item. Significant levels of missing data were fairly small with just 213 respondents missing or making an error on more than 10 items.

The characteristics of those with higher levels of item non-response can again be examined. Although men were found above to be less likely to complete the self-completion questionnaire, Table 5 shows that amongst responders, high item non-response (defined as 5 or more missing items) was more prevalent among women. Bivariate analyses, based on these 379 high item non-responders suggest that high item non-response is negatively associated with education (high levels of non-response was most prevalent amongst those with no qualifications) and the presence of a partner. Using the Chi-squared test, all these bivariate associations were found to be significant at the $p<0.05$ level.

Table 5: High levels of item non-response (5 or more items missed) by characteristics of cohort members completing the self-completion questionnaire

Characteristic	Proportion with high levels of item non-response	Total N
Male	0.036	151
Female	0.050	228
No spouse or partner	0.068	121
Lives with spouse or partner	0.037	256
No qualifications	0.118	109
Quals at NVQ level 1	0.054	53
Quals at NVQ level 2	0.040	89
Quals at NVQ level 3	0.034	52
Quals at NVQ level 4	0.024	67
Quals at NVQ level 5 or 6	0.020	8
All	**0.043**	**379**

The paper questionnaire also included an open question at the end which asked respondents to imagine that they were 60-years-old and to write a few lines about their lives (their interests, home life, health and wellbeing, and employment). Answering an open-question, particularly one of the this nature, requires considerably more cognitive effort than answering the closed questions. As such, as one might expect the level of non-response on this item was considerably higher than any of the closed items (16 per cent compared with the highest level of non-response to a closed item which was 2 per cent). However, unlike overall levels of response to the questionnaire and high levels of item non-response, bivariate analyses revealed no significant associations between levels of response to the open question and sex, presence of partner, or qualifications.

The use of more than one mode of data collection in 2008 sweep of the 1958 birth cohort study provides an opportunity to examine whether, as has been found in other studies, *item* non-response was higher in the self-completion questionnaire than in the face to face interview. Clearly to test this properly it would be necessary to have the same battery of questions repeated in each mode, or to have an experimental cross-over design such that a random sample of cohort members answered a battery of questions as part of the paper questionnaire, while another random group answered the same questions as part of the face to face interview. However, given that the survey was designed to collect information from cohort members as efficiently as possible, rather than being a methodological experiment, this type of experimental cross-over was not implemented. It is still possible, however, to gain some insights by comparing the item non-response rates for a battery of questions

asked in the face-to-face interview and a similar battery of questions in the self-completion paper questionnaire.

A series of nine questions, which make up a shortened version of the malaise inventory, were asked as part of the face to face interview. These have two response categories ('Yes' and 'No') so that cohort members indicate whether they experience each symptom or not. As part of the paper questionnaire there were a series of seven questions which make up two of the sub-scales of the SF36. These also focus on health symptoms and have two response categories ('Yes' and 'No'). As is summarised below in Table 6, analysis showed that there was considerably more item non-response in the context of the paper questionnaire than in the context of the face to face interview. Indeed it is interesting to note how low item non-response is to the questions in the malaise inventory, even though these were presented to individuals in the context of the self-completion computer assisted section of the interview. This would perhaps indicate that it is the presence of the interviewer that is encouraging response rather than cognitive burden or literacy problems increasing non-response.

Table 6: Item response in CASI and PAPI

	PAPI mode		**CASI mode**	
	Seven SF36 questions		**Nine malaise questions**	
	N	**%**	**N**	**%**
Number completing all questions	8703	99.03%	9619	99.82%
Number missing at least one question	85	0.97%	17	0.18%
Total	8788		9636	

Conclusions

In conclusion therefore, as we would expect, those who did not complete and return the paper questionnaire were not a random subgroup of those who responded to the face to face interview at age 50. Those who did not complete the questionnaire are more likely to be male, with low qualifications, living without a spouse or partner and with poor self-reported health. However, even among these groups, response rates to the self-completion questionnaire were over 80 per cent. Researchers should be aware that including variables from the self-completion questionnaire in their models will reduce the sample size by around 10 per cent and the non self-completers are unlikely to be 'missing at random'[iii]. The response rate to the paper questionnaire was however very high, in comparison with what one would expect in response to a paper questionnaire, and as discussed elsewhere (Brown 2008) this was a very cost-effective method of collecting data in the context of a long-running cohort study.

References

Bowling, A. (2005) Mode of questionnaire administration can have serious effects on data quality. Journal of Public Health, 27, 3 pp.281-291. doi:10.1093/pubmed/fdi031.

Brazier JE, Harper R, Jones N, et al. (1992) Validating the SF-36 health survey questionnaire: new outcome measure for primary care. Br Med J 1992; 305: 160–164.

Brown, M. (2008) An evaluation of the paper self completion questionnaire planned for use as part of 2008 sweep of the National Child Development Study. CLS Technical Report. London: Centre for Longitudinal Studies.

Brown, M. & Dodgeon, B. (2010) NCDS Cognitive Assessments at Age 50: Initial Results. CLS Working Paper 2010/1. London: Centre for Longitudinal Studies.

Carpenter, J. & Plewis, I. (2011). Analysing longitudinal studies with non-response: issues and statistical methods. Chapter 23 (pp. 498-253). In Williams, M. & Vogt, P. (Eds.) Handbook of Methodological Innovations, Newbury Park, Ca: Sage.

De Leeuw ED, van der Zouwen J. (1988) Data quality in telephone and face-to-face surveys: a comparative meta-analysis. In: Groves RM, Biemer PP, Lyberg LE et al. eds. Telephone survey methodology. New York: John Wiley and Sons, 1988.

Dodgeon, B. Hancock, M. Johnson J. Parsons, S. (2011) Deriving Highest Qualification in NCDS and BCS70. CLS Data Note 2011/1. London: Centre for Longitudinal Studies.

Goldberg, L. R. (1999). A broad-bandwidth, public domain, personality inventory measuring the lower level facets of several five-factor models. In I. Mervielde, I. Deary, F. De Fruyt, & F. Ostendorf (Eds.),Personality Psychology in Europe, Vol. 7 (pp. 7-28). Tilburg, The Netherlands: Tilburg University Press.

Harris LE, Weinberger M, Tierney WM. (1997) Assessing inner-city patients' hospital experiences. A controlled trial of telephone interviews versus mailed surveys. Med Care 35: 70–76.

Nicholaas G, Thomson K, Lynn P. (2000)The feasibility of conducting electoral surveys in the UK by telephone. Centre for Research into Elections and Social Trends. London: National Centre for Social Research, and Department of Sociology, University of Oxford.

Tourangeau R. (1984) Cognitive Sciences and survey method. In Jabine T, Straf M, Tanur J, Tourangeau R, eds. Cognitive aspects of survey methodology: building a bridge between disciplines. Washington DC: National Academy Press.

[i] The variable describing cohort members' highest level of qualifications was derived using data from all relevant sweeps of the longitudinal study. For further details see the data note on highest level of qualifications by Dodgeon et al (2011).

[ii] The association here was not significant

[iii] Readers who are interested in tackling item non-response and the issue of 'non-ignorable' non-response should consult Carpenter and Plewis (2011).