ATTRITION BIAS IN TELEPANEL RESEARCH

Juan Felix & Dirk Sikkel [1]

Abstract

CentER*data* operates a telepanel, in which the respondents fill in a questionnaire each week on their computer. The panel now is seven years old. This raises the question how long respondents may be allowed to stay a memberof the panel. Part of this question can be answered by examining the attrition bias. If 'old' panel members differ systematically from 'new' members with respect to variables that can not be controlled in the selection process this would be a reason to limit panel membership. The analysis shows that there are no such differences with respect to a large range of psychological variables.

1. Introduction

Nowadays, household panels are available in a multitude of varieties. In official statistics and social sciences most of the household panels take part in a research project once a year. In consumer research, panels usually are approached more often to deliver data to the research agency, e.g. in scanner panels, TV panels and expenditure panels. Even when data are collected in an automated way, e.g. with a scanner, the question arises whether the respondents who stay in the panel for a long time are an atypical selection of the target population they are supposed to represent. This question becomes more pressing in a heavily used panel which often participates in opinion research, sociological research, psychological research etcetera. These latter types of research are predominant in a telepanel. Such a panel consists of a random sample of households who are provided with a PC and a modem at home. The panel members receive a questionnaire which is transmitted by modem and phone. The answers to the questions are sent back to a central computer using the same procedure of data transmission. CentERdata, a research bureau at the Tilburg University, uses a telepanel which consists of approximately 2000 households. In principle, every member of a selected household becomes a member of the telepanel. Once these household members agree to become panel members, they are expected to answer survey questions on a weekly basis. The panel members fill in the questionnaires during the weekend. Each year, the panel members are asked to answer up to 50 questionnaires. It will be clear that this puts a great burden on these panel members, considering that their participation in the telepanel is voluntary. Despite the burdensome task of completing up to 50 questionnaires a year, the duration of panel membership often exceeds several years.

New households are canvassed in different stages. In the first stage, new panel members are selected by telephone interviews (CATI). The sample for these interviews is based on a random selection of telephone numbers of Dutch private households. In the interview some

---

[1] CentER*data*, P.O.Box 90153, 5000 LE Tilburg, The Netherlands.
Felix: phone: +3113-4668326; fax: +3113-4662764; e-mail: A.F.M.Felix@kub.nl
Sikkel: phone: +3113-4668323; fax: +3113-4662764; e-mail: Dirk.Sikkel@kub.nl

basic questions are answered by the head of the household. About 40% of the respondents refuse to answer any questions. At the end of the interview, the respondents are asked whether they are prepared to participate in panel research. About 40% of the respondents are absolutely not prepared to participate in the panel. The other 60% of the respondents are considered to be potential panel members. These potential panel members are registered and stored in a basic 'sign up' file. Every week new households are selected from this basic file. The chance of being selected for the panel depends on the distribution of six demographics in the panel (see Appendix). The households who are selected from this basic 'sign up' file receive an introduction letter with some basic information on panel research. These households are interviewed for a second time. In this interview, the head of the household has to decide whether or not some or all members of the household will actually become a member of the panel.

Table 1. Duration of stay in the telepanel by quartile.

	weeks
first quartile	0 - 75
second quartile	76 – 141
third quartile	142 – 216
fourth quartile	217 – 385

At present there is no limit to the length of a period during which a member may stay in the panel. This leads to the distribution of stay in the telepanel as given by table 1. About 25% of the panel members have been in the panel for more than four years. Clients of CentER*data* often express their disbelief that this is a sound situation. In this paper we will investigate if there are reasons to limit the length of panel membership and, if this is the case, what the optimum length should be. We start in section 2 with an overview of all possible sources of bias in a telepanel; in section 3 we discuss possible bias with respect to demographic variables; section 4 deals with psychological variables. Section 5 concludes.

2. Bias in a telepanel

In principle, every research organisation has to deal with the problem of nonresponse. When nonresponse occurs, this means we have to deal with missing observations. Missing observations may bias the estimations of population characteristics. Nonresponse may occur in different stages of the process:

1. Initial nonresponse. This is the nonresponse that occurs in the process that precedes the membership of the panel, the recruitment stage as described in section 1.
2. Wave nonresponse. This occurs when panel members do not answer survey questions during one or more waves but participate in subsequent waves (see also Verbeek 1991).
3. Item nonresponse.This occurs when one or more items are not completed for a unit that otherwise provides responses (Lepkowski 1989). So in a panel situation this is non response to a single item within one wave. In a telepanel this is rather uncommon.
4. Attrition. This is the situation when a panel member leaves the panel and has to be replaced by a new member. It can be considered as permanent wave nonresponse.

These types of nonresponse may or may not cause bias, or, more likely, they may cause bias with respect to some research areas. This depends on the differences between the respondents

and the non respondents with respect to the target variables and the purpose of the studies. Luijten and Hulsebos (1997) stated that it is not clear whether the initial nonresponse plays a significant role in the permanent proces of keeping the panel at strength. When it is the purpose of a project to study trends, initial nonresponse may be irrelevant, even when target variables are biased. Apart from the different types of nonresponse, there may be another source of bias.

5. Panel education. The possibility exists that panel members learn by filling in questionnaires. They may become interested in political issues or consumer issues by the panel process itself. Moreover, they are trained in filling in questionnaires (test-wiseness) which may cause their response behaviour to be atypical.

In this paper we will restrict ourselves to attrition. The other possible sources of bias are also under investigation but require more time because their experimental setups are longitudinal.

One of the major concerns for an organisation that uses a telepanel is keeping the attrition rates as low as possible. Besides the high costs of replacing panel members that exit the telepanel by new ones, attrition in itself does not have to be a problem. Attrition becomes a problem if it is selective. Taris has illustrated the effect of selective nonresponse in longitudinal research (Taris 1995). The cumulation of selective nonresponse results in a sample that is not representative over time.

Selective attrition will lead to underrepresentation of certain subpopulations in the panel. This underrepresentation can occur with respect to two kinds of variables:

1. variables for which the population distribution is known. So, correction is possible by poststratification if we measure these variables on entry into the panel.

2. variables for which the population distribution is not known; this is, of course, the majority of variables. Here we focus on psychological traits, as they may be considered as causal variables for attrition when it is assumed that a certain personality structure causes individuals to continue their panel membership.

A number of 'procedures' are designed to decrease attrition in the panel. First, CentERdata is honest about the respondents burden when they participate for a longer period of time. Respondents have to know what to expect when they become a member of the panel. Second, households need a computer and modem to answer the survey questions. Households who don't have a PC at home are provided with one by CentERdata. Third, the respondents are motivated by feedback on the surveys they participate in. Although in some projects the respondents have to be kept in the dark about purpose and results, information about results of public projects is much appreciated. Fourth, CentERdata uses a helpdesk to assist households with certain technical problems, to answer possible questions about the survey, or to register any reason why panel members are not able to participate for a short or longer period of time.

We have limited ourselves mainly to examining attrition with respect to demographic and psychological variables because they are not influenced by the participation in the panel. In other words, participation in the panel does not cause changes in the value of these variables. So if we find results that indicate differences in the personality structure between panel

members with a long duration in the panel and relatively new members these differences can be attributed to selective attrition. The central question is:

Do we have to limit the duration of panel membership?

This leads to the following sub-questions:

a) *In what respect are respondents who participate in the panel for a long time different from respondents who only participate for a short time?*

b) *If there are any differences, can we make some corrections in the replacement procedure of old panel members by new ones?*

If the results of our investigations indicate a relation between socio-demographics and duration of panel membership we can adapt the selection procedure of new panel members on the basis of demographic variables like urbanisation, composition of household, age, income and political preference, such that a correction is made for attrition. If, however, there is a relation between socio-psychological characteristics and the time in the panel, it is not possible to make corrections as (1) it is very costly to measure psychological traits in the selection interview and (2) we do not know the distribution of these psychological characteristics in the population.

3. Demographic variables

The first part of the analysis concerns the basic demographic variables. In the analyses we focused on the panel membership of individual household members. We did not examine the panel duration of complete households because the duration of panel membership can differ between household members of the same household. These differences can partly be explained by changes in the composition of households and partly by differences in age, like in the following examples. Young household members may leave the household if they are going to study. Or a household member may die while his wife or (her husband) continues to participate in the panel. Two people may decide to live together and one of them joins the panel whilst the other is already a member. These are all examples of a 'natural' increase or decrease of the panel.

In this study we examined the panel membership duration of individual panel members who are 18 years of age and above. These household members are selected for most or all of the questionnaires. This implies that their panel membership can be more burdensome than that of the younger household members. On the other hand the older household members may get a strong sense of belonging to the panel. Furthermore, we assume that the decision to quit the panel is mostly not taken by a child in the household.

First, we examined the attrition with respect to the demographics that are used in the selection procedure of new households. These demographics are composition of household, income, age, political preference and urbanisation. In the subsequent analyses we did focus on demographics that are not used in the selection procedure but may influence the duration of panel membership. These are education, most important occupation, family size, living with a partner (or not), gender, region and degree of rurality (ranging from urban to rural).

We have calculated a measure of association between 11 of these demographical variables and time of panel membership. All of these demographical variables are categorical. Age of panel members is measured as a numeric variable. Therefore we have calculated the correlation coefficient of age with time of panel membership.

The highest correlation including duration of membership was with the variable age (r=0.251 , which indicates that older panel members stay longer in the panel than younger panel members). The influence of age is also reflected by the average panel time of some specific 'old age' categories. The older panel members live relative often alone or with a partner but without children, they are often pensioners and they have a relative low income. Our results indicate that the average panel membership of these three categories is relative high.

Table 2 : Association: demographics with duration of panel membership, with correction for age (n=3834).

Demographics	η (eta)	β (beta|age)	β (beta|age and income)
in selection procedure			
Urbanisation	0.07[c]	0.07[c]	0.06[c]
Income	0.21[a]	0.16[a]	-
Composition of household	0.19[a]	0.10[a]	0.08[a]
Voting behaviour national election	0.06[d]	0.04[d]	0.03[d]
not in selection procedure			
Education	0.12[a]	0.09[a]	0.09[a]
Partner (Yes/No)	0.04[c]	0.02[d]	0.01[d]
Family size	0.18[a]	0.09[a]	0.08[a]
Gender	0.02[d]	0.01[d]	0.02[d]
Most important activity	0.20[a]	0.10[a]	0.10[a]
Region	0.09[a]	0.08[a]	0.08[a]
Degree of rurality	0.06[c]	0.05[c]	0.04[c]

[a] F-ratio is significant at 0.001 level; [b] F-ratio is significant at 0.01 level; [c] F-ratio is significant at 0.05 level; [d] F-ratio is not significant.

The results for the categorical variables are based on ANOVA (analysis of variance). The first column of table 2 shows low to moderate correlation ratios (eta). Eta squared can be interpreted as the proportion of the total variance in the dependent variable (mean time of panel membership) that can be accounted for by knowing the values of the independent variables (the ten demographics). The results of the first column of table 2 indicate income and most important activity have some relationship with the duration of panel membership. Since most variables in table 2 vary with age, this raises the question if age is not the predominant factor in all relations.

This question can be answered by calculating the second column of table 2, the betas, which are the correlation ratios adjusted for the influence of age. Now we see that, as expected, all relations decrease. Given that age is one of the variables in the selection procedure, controlling for age (i.e. replace leaving panel members by new members such that the distribution by age in the panel resembles the distribution by age in the population) reduces the demographic attrition bias considerably. Fortunately, the variables which are not used in the selection procedure also have, given age, hardly a relationship with panel membership duration. Still it is also necessary to check on income to guarantee that attrition does not

cause bias with respect to this variable. Most of the relations between the demographics and panel time decrease even more if we also correct for income, (table 2, last column).

4. Psychological variables

The psychological variables under investigation can be divided into three categories: general, values and personality. For all of these categories hypotheses were formulated that reflect lines of reasoning that are often heard in practice, although there is no scientific evidence to support them. All hypotheses were tested with the use of well established scales. Before we show the results we first give some comments on the hypotheses.

Need for cognition (NFC) represents the tendency for individuals to engage in and enjoy thinking (Cacioppo and Petty 1982). The original scale is comprised of 34 items each scored –4 to +4 . The categories range from very strongly agree (+4) to very strongly disagree (-4). (Cronbach's $\alpha = 0.89$). It's plausible to apply the need for cognition to panel membership since panel members are asked to think about diverse subjects when they answer survey questions (hypothesis 1).

Innovativeness or the openness of information processing is assumed to be a personality trait underlying the adoption of innovations. Innovators are described as individuals open to new experience and novel stimuli, as possessing the ability to transform information about new concepts, ideas, products or services for their own use, and as having a low threshold for recognizing the potentional application of new ideas. (Leavitt and Walton 1975, 1988). The measure for innovativeness consists of 24 items. Each statement is evaluated in terms of strong agreement to strong disagreement. (Cronbach's $\alpha = 0.84$). At first, participating in panel research may be a new and different thing for panel members. But, in the long run, panel members may get bored by this type of research (hypothesis 2).

The Social Desirability Bias (SDB) measures the degree to which people describe themselves in socially acceptable terms to gain the approval of others (Crowne and Marlowe 1960). The scale is comprised of 33 statements that are answered in terms of true or false. (Cronbach's $\alpha = 0.64$). People with higher scores on the scale are more inclined to respond to questions in a manner that they deem socially desirable. Panel members who give socially desirable answers may find it difficult to say that they want to stop their participation (hypothesis 3).

Loneliness involves situations in which the number of achieved relationships is smaller than desired, or when the existing relationships fail to attain the desired degree of intimacy. The type of deprivation refers to the nature and intensity of the missing relationships (de Jong-Gierveld and Kamphuis 1985). We used a unidimensional scale with 11 items to measure loneliness. These items all relate to one latent characteristic: a type of deprivation. (Cronbach's $\alpha = 0.80$). The assumption underlying the hypothesis is that the partictipation to the panel serves as a substitute for social contact. CentER*data* has the role of a friend. (hypothesis 4).

Nine hypotheses are based on the List of Values (LOV), which comprises 9 values that are central to people in living their lives, particularly the values of life's major roles (i.e. marriage, parenting, work, leisure and daily consumptions). Positive aspects of panel membership could be the satisfaction of the need to belong to a group, warm relationships with others, the feeling that one is well respected, to have a sense of security, self-respect, to

have a sense of accomplishment and self-fulfillment (hypotheses 5, 7, 8, 9, 11, 12 and 13). Panel membership could be negatively valued because it may be not exciting or enjoyable (hypotheses 6 and 10).

Exhibit 1: Hypotheses with respect to psychological variables and duration of panel membership

General

Hypothesis 1: *Panel members who are relatively high in need for cognition are more likely to stay in the panel for a long time.*

Hypothesis 2: *Panel members who are always open to new experience and stimuli will quit sooner than panel members who are not constantly looking for new experience.*

Hypothesis 3: *Panel members who are more inclined to give socially desirable answers will have a higher duration of panel membership.*

Hypothesis 4 : *Panel members who feel very lonely will have a higher duration of panel membership.*

Values

Hypothesis 5 : *Panel members who highly value a sense of belonging will stay long in the panel.*

Hypothesis 6 : *Panel members who find it important to have excitement in their lives will have a low duration of panel time.*

Hypothesis 7 : *Panel members who highly value warm relationships with others will have a high duration of panel membership.*

Hypothesis 8 : *Panel members who think of self-fulfillment as being very important will stay longer in the panel.*

Hypothesis 9 : *Panel members who find it important to be well respected will stay longer in the panel.*

Hypothesis 10 : *Panel members who highly value fun and enjoyment of life will have a low duration of panel time.*

Hypothesis 11 : *Panel members who think of security as being very important will stay longer in the panel.*

Hypothesis 12 : *Panel members who find it very important to have self-respect will stay longer in the panel.*

Hypothesis 13 : *Panel members who highly value to have a sense of accomplishment will have a high duration of panel membership.*

Personality

Hypothesis 14: *Panel members who are more emotionally stable will have a higher duration in the panel than members who are less emotionally stable.*

Hypothesis 15: *Panel members who are very extravert will stay longer in the panel than members who are less extravert.*

Hypothesis 16: *Panel members who have a high sense of agreeableness will stay longer in the panel.*

Hypothesis 17: *Panel members who are very conscientious are more likely to stay in the panel than panel members who are not so conscientious.*

Hypothesis 18: *Panel members who are less autonomous will have a higher duration in the panel compared with members who are more autonomous.*

Furthermore, the panel members have judged their own personality on the basis of 100 single trait terms. In her doctoral thesis Jolijn Hendriks uses these 100 items to construct a five-factor personality inventory (FFPI). "(…) the five-factor personality inventory assesses five broad dimensions (factors) of individual differences in behaviour: Extraversion,

Agreeableness, Conscientiousness, Emotional Stability, and Autonomy" (Hendriks, 1997). Estimates of internal consistency reliability for these five factors are respectively 0.85, 0.82, 0.83, 0.83 and 0.77 . As mentioned earlier, complete households or single household members participate in the panel for several years. One could assume that such an achievement must imply that these panel members have special personal qualities. Such as doing things in an orderly and disciplined way, being inclined to help others, being open to others and thinking positive (hypothesis 14 to 18).

Table 4 Correlation psychological characteristics with duration of panel membership.

Psychological characteristics	r	n
General		
loneliness	-0.010	1914
social desirability	0.036	1714
need for cognition	-0.071	1630
innovativeness	-0.032	1638
List of values		
sense of belonging	-0.004	2264
excitement	-0.043	2264
warm relationships	-0.045	2264
self development	-0.075	2264
being well respected	-0.062	2264
enjoying	-0.048	2264
security	-0.042	2264
self respect	-0.063	2264
achieving something	-0.083	2264
Personality traits		
Emotional Stability	0.017	2494
Extraversion	-0.025	2494
Agreeableness	-0.005	2494
Conscientiousness	0.049	2494
Autonomy	0.024	2494

Table 4 shows that there is hardly any relation between the various psychological characteristics and duration of panel membership. For example, panel members who stay longer in the panel do not feel more lonely than panel members who stay for a short period of time. We have calculated partial correlation coefficients between the psychological characteristics and attrition to control for a possible 'suppressing' effect of age. For example, if there is a relation between age and loneliness this could fade away the relation between loneliness and attrition. The results indicate however that there is no such effect in the sample data. We may conclude that there is no convincing evidence that supports even one of the 18 hypotheses.

5. Conclusion

It appears to be highly unlikely that panel members who stay in the panel for a long time are freaks, loners, idealists, neurotics or belong to any other psychological minority which makes them suspect as respondents for survey research. On the basis of our research results we did not find support for any of the hypothesis that were formulated with respect to relations with duration of panel membership stated here.

Our results indicate that panel membership is related to age of the individual household members. Older panel members are more likely to stay in the panel than younger panel members. This implies that in the recruitment stage of new panel members we have to make certain that the age distribution is correct. In the selection procedure of new panel members age is one of the conditional variables to add new members to the panel. Given age, duration of panel membership is hardly associated with demographics like education, family size, composition of household, living with a partner or not, urbanisation, region and degree of rurality. Only income has a considerable relation with membership duration, given age. Since this variable is considered in the recruitment process, this does not have to lead to attrition bias.

Many variables have been taken into account in this study. On the basis of the evidence presented here
we may conclude that there is no reason to put a maximum to the panel membership duration. The question arises if the evidence is conclusive. So far, one essential element has been left out: the cognitive training of the respondent which may increase his knowledge about the world and his ability to fill in questionnaires. In order to obtain conclusive answers about this aspect 'old' and 'new' respondents have to be compared over time. As a consequence, it will take still one or two years to find these answers.

References

Cacioppo, John T. and Richard E. Petty, 1982, The Need for Cognition, Journal of Personality and Social Psychology, vol 42, no. 1, pp. 116-131.

Crowne, Douglas P. and David Marlowe, 1960, A New Scale of Social Desirability Independent of Psychopathology, Journal of Consulting Psychology, vol. 24, pp. 349-354.

Hendriks, J, 1997, The construction of the Five-Factor Personality Inventory (FFPI), Ph. D. Thesis, University of Groningen.

de Jong-Gierveld, J. and F. Kamphuis, 1985, The Development of a Rasch-Type Loneliness Scale, Applied Psychological Measurement, vol. 9, pp. 289-99.

Kahle, Lynn R., 1983, Social Values and Social Change: Adaptation to Life in America. New York: Praeger Publishers.

Leavitt, Clark and John Walton, 1975, Development of a Scale for Innovativeness; in Mary Jane Schlinger (ed.), Advances in Consumer Research, Vol. 2, Ann Arbor, MI: Association for Consumer Research, pp. 545-554.

Lepkowski, J.M., 1989, Treatment of Wave Nonresponse in Panel surveys; in D. Kasprzyk (ed), Panel surveys, New York, John Wiley.

Luijten A.L.J.M. and W.H.L. Hulsebos, 1997, "Panelbias: een fenomeen in continu onderzoek" in Recente ontwikkelingen in het marktonderzoek, Jaarboek 1997 van de Nederlandse Vereniging voor Marktonderzoek en Informatiemanagement, De Vrieseborch, Haarlem, pp. 61-75.

Taris, T. W., 1995, Nonrespons en representativiteit in panel studies: Een korte notitie. Kwantitatieve Methoden, 48, 41-44.

Verbeek, M.J.C.M., 1991, The design of panel surveys and the treatment of missing observations, Ph. D. Thesis, University of Tilburg.

Appendix. The recruitment of panel members

New panel members are selected by telephone interviews (CATI). In this first interview some, basic questions are asked. These questions are:

- newspaper(s) read regularly
- political preference (party voted for during last national elections)
- age
- composition of household
- income.

Every household that is willing to participate in the telepanel has a chance of being selected to become a new member. Households that are willing to cooperate in the panel are registered and stored in a basic 'sign up' file. Every week new households are selected from this basic file. The chance of being selected for the panel depends on the following variables

Household characteristics:
- urbanisation
- composition of household

Head of household characteristics:
- income
- age
- political preference.

In the ideal situation the distribution of these six demographics in the panel is exactly representative of the distribution in the population. The new panel members are drawn from the sign-up file in such a way that the distribution within the panel matches the population distribution as close as possible.

Ontvangen: 12-05-1998