Transportation (2009) 36:47–64
DOI 10.1007/s11116-008-9185-5

# Estimating the robustness of questionnaire results: lessons from a mixed-mode survey of expectations for tele-working and road-based business travel

**Peter Bonsall · Jeremy Shires**

Published online: 25 December 2008
© Springer Science+Business Media, LLC. 2008

**Abstract** The robustness of questionnaire results to various forms of bias are explored in the context of a dual-mode (web and hardcopy) survey of employers' anticipations of levels of employee commuting and business travel activity under a range of future ICT scenarios. The questionnaire incorporated several innovative features which, together with the dual-mode format, allowed an unusually wide range of analyses. For example: the robustness of respondents' opinions was tested by examining the effect of incorporating alternative versions of a briefing text, one being very positive and one very negative, about the role of ICT; instrument bias was identified via detailed comparison of the results from the two versions of the questionnaire; and the impact of exogenous factors which are often ignored or taken as constant was assessed via special supplementary questions. Analysis showed that the robustness of opinions and expectations varied and was influenced by respondent characteristics, and that results from the two versions of the questionnaire differed significantly. It is concluded that opinions and expectations are less robust, and questionnaire results are more subject to bias and myopic interpretation, than is generally recognised and that web-based surveys seem particularly vulnerable to sampling bias. Methods are suggested for measuring robustness, for reducing bias and for validating and contextualising results. The use of contrasting briefing texts is recommended as a means of establishing the robustness of opinions and expectations while supplementary questions are recommended for validating and contextualising SP and SE exercises.

**Keywords** Robustness · Instrument bias · Response bias · Framing effects · Contrasting briefings

## Introduction

The analysis reported in this article began in the context of a multinational project ("POET"), funded by the European Union, which sought to explore the impacts of the

P. Bonsall (✉) · J. Shires
Institute for Transport Studies, University of Leeds, Leeds LS2 9JT, UK
e-mail: p.w.bonsall@its.leeds.ac.uk

increasing availability of information and communication technologies (ICT) on transport and travel demand. The overall conclusions from the project are presented elsewhere (see for example POET 2005a, b; De Jong et al. 2006).

This article reports on our further work on a number of methodological issues which arose in the course of the POET project and which provide useful insights into alternative approaches to the collection of behavioural data and preferences. This further work capitalised on the existence of data from two versions of the same questionnaire—an interactive version placed on a website and a hard-copy version distributed via conventional mail—and on the fact that we had incorporated a number of features to test the robustness of the results obtained. The questionnaire included revealed preference data (on current amounts of working from home and of road-based business travel), stated expectation data (on expected occurrences of these activities in each of a range of future scenarios), and respondents' ratings of the influence of each of a series of factors on their expectation of future outcomes. A particularly innovative feature was the use of randomly allocated briefing texts to test the resilience of respondents' opinions.

Analysis of the resulting data allows us to explore and comment on a number of important issues affecting the design of questionnaires and, more broadly, the interpretation of data collected using questionnaires. One of these issues is s*urvey mode bias* (a type of instrument bias, whereby different results to emerge when the same questions are posed via different media—in this case via the web or hardcopy—see for example Dillman 2000; Kaplowitz et al. 2004; Arentze et al. 2005). Another is *non-response bias* (resulting from the fact that some people are more/less likely to respond to a questionnaire on a given topic presented in a given manner—see for example Sax et al. 2003). Another is the *framing effect* (whereby the responses to specific questions are influenced by the context in which they have been set—see Tversky and Kahneman 1981). Finally there is the risk of what we term *myopic interpretation*—which is introduced when key factors are ignored in the specification of scenarios in Stated Preference exercises.

## Data source

### Questionnaire design and implementation

The questionnaire was designed to provide data with which to calibrate models of factors affecting employers' expectations of future levels of home-working and road-based business travel. It had been decided that all the POET questionnaires (ours being one of four) would be web-based because of the anticipated cost savings (reduced printing, postage and data-correction/processing costs), because the presentation of scenarios was expected to need to be individually tailored and randomised (which is easily achieved in an on-line questionnaire) and because the subject matter (ICT futures) was thought particularly amenable to a web-based presentation.

The structure and content of the questionnaire is summarised in Table 1.

The background information (A) was required in order to allow investigation of factors influencing future expectations and to allow differential weighting of the results in the event that the profile of the responding population did not match that for which forecasts were required (not only was a degree of non-response bias inevitable with this kind of survey but we anyway wanted to make forecasts for different populations).

The briefing texts (B1 and C1) were an innovative aspect of our questionnaire. They were included because, believing that peoples' expectations of the tele-working, as

**Table 1** Questionnaire structure and content (web version)

| |
|---|
| A: Background information on the organisation and respondent: |
| 1. Respondent's role in organisation or group? (e.g. senior manager, middle manager) |
| 2. Size of company/group |
| 3. Nature of business? (e.g. retailing, manufacturing, corporate services) |
| 4. Location of premises? (e.g. city centre, business park) |
| 5. Policy on employees working from home? (e.g. encourage, discourage) |
| 6. Current extent of working from home? (% of workforce on a typical day) |
| 7. Does respondent work from home? |
| 8. Access to video-conferencing facilities? (e.g. in building, on site, none) |
| 9. Does respondent use videoconferencing equipment? |
| 10. Current extent of business travel? (number and length of trips per head per year) |
| 11. Expected size of organisation or group (or its successor) in 2010? |
| B: Working from home |
| 1. Briefing text on working from home (randomly positive or negative) |
| 2. Assessment of briefing text? |
| 3. Expected level of working from home in 2010? (for each of 6 scenarios) |
| C: Business travel (road based—excluding deliveries) |
| 1. Briefing text on business travel (randomly positive or negative) |
| 2. Assessment of briefing text? |
| 3. Expected volume of business trips travel in 2010? (for each of 5 scenarios) |
| 4. Expected length of business trips in 2010 (for each of 2 scenarios) |
| D: Influences on expectation |
| 1. Impact of specified factor on level of home working in 2010 (for each of 6 factors) |
| 2. Impact of specified factor on level of business travel in 2010 (for each of 6 factors) |

reported in previous studies, had perhaps been overly influenced by media hype, we wanted to test the robustness of respondents' opinions to outside influences. Our chosen method was to expose them, just prior to asking them for their expectations, to texts which purported to represent third party opinions. The texts seen by a given respondent were either generally positive or generally negative about tele-working (at B1) and similarly either generally positive or generally negative about the extent to which business travel might be replaced by ICT (at C1). The allocation of generally positive or negative texts at points B1 and C1 was purely random. Questions B2 and C2 asked how balanced the respondent thought that the text was and were designed to indicate the distribution of opinion within the responding population.

The scenarios presented to respondents in questions B3, C3 and C4 are listed in Table 2 and reflect the range of "futures" that were of interest to the POET project. However, since they did not encompass the full range of factors that might influence future expectations, respondents were asked (at D1 and D2) to indicate the amount of influence that specified factors (e.g. the cost of home office equipment and the speed of broadband links) had on their expectations.

As noted above, the original questionnaire was designed for on-line implementation. However, as will be discussed below, the response to the web version of the questionnaire was disappointing and we decided to revert to a hard copy version when the opportunity arose to extend the sample and to analyse the resulting data in more depth. A mixed mode

**Table 2** The scenarios presented to respondents

*Six scenarios for questions on home-working in 2010:*

(1) The cost of business office space and the cost of acquiring and using state-of-the-art home office equipment are similar to 2004

(2) The cost of business office space have risen by 25% but the cost of acquiring and using state-of-the-art home office equipment is similar to 2004

(3) The cost of business office space is similar to 2004 but the cost of acquiring and using state-of-the-art home office equipment has fallen significantly

(4) The cost of business office space have risen by 25% but the cost of acquiring and using state-of-the-art home office equipment have fallen significantly

(5) As (1) but the travel conditions faced by commuters have deteriorated such that journeys take 20% longer

(6) As (4) but the travel conditions faced by commuters have deteriorated such that journeys take 20% longer

*Five scenarios for questions on the number of business-related trips in 2010:*

(1) The capability and price of ICT remain as they are in 2004

(2) The capability of ICT continues to increase at the current rate and its price continuous to fall at the current rate

(3) The capability of ICT increases in line with the most bullish forecasts and its price falls significantly

(4) As (1) but road travel gets worse (journeys take 20% longer)

(5) As (3) and road transport gets worse (journeys take 20% longer)

*Two scenarios for questions on length of business-related trips in 2010:*

(1) The capability and price of ICT, and the travel conditions, remain as they are in 2004

(2) The capability of ICT increases in line with the most bullish forecasts, its price falls significantly, and a typical journey takes 20% longer

approach of this type is commonly adopted by market research firms because it offers the opportunity to improve response rates and circumvent many of the access barriers (ex-directory telephone numbers, gated communities, multiple email addresses, spam filters) that exist with individual survey methods (Schaefer and Dillman 1998; Dillman 2000).

The structure and content of the hard copy version (which can be viewed at www.its.leeds.ac.uk/poet/paperquestionnaire) was almost identical to that of the web version except that, in order to meet page layout constraints, the question on current levels of business travel (A10 in Table 1) was placed at the beginning of Section C (before the briefing text), and the two questions in Section D were brought forward to the end of sections B and C, respectively.

## Sampling and response rates

The target population for the web questionnaire was managers of organisations employing 10 or more people in the UK, Netherlands, Greece, Italy and Israel (this selection of countries being dictated by the membership of the POET consortium). Recruitment occurred during the second half of 2004 and, in line with best practice as recommended by Dillman (2000), used a variety of media; initial contacts were made via phone calls, letters and emails, and phone calls were used to chase up a selection of non-respondents. The majority of the initial contacts were randomly selected from trade directories (e.g. Dun and Bradstreet and the FAME database) but this sample was supplemented by personal contacts.

Around 500 telephone calls were made to personnel management staff to explain the background to the project and request participation. Respondents were offered a discount voucher (entitling them to a 25% discount on meals at certain restaurants) and those who agreed to participate were sent an email providing the web address. This exercise produced fewer than 90 completed questionnaires and was very expensive to mount. An alternative, cheaper, approach involved letters or emails introducing the project, offering the opportunity to win a case of champagne and access to the project results, and inviting respondents to log on to the website (in the case of the emails this was via a hotlink). Around 200 letters were sent to Directors of Human Resources but this produced only a handful of completed questionnaires. The emailing was on an altogether larger scale; several thousand emails were dispatched (some direct to Human Resources managers, and some via the organisation's "info@" or "postmaster" addresses with a request to forward them to the organisation's director of Human Resources) but this generated only slightly over 100 completed questionnaires.

Follow-up phone calls and investigation of the site log revealed that we had suffered from many of the problems which can beset web-based surveys (see for example Smith 1997; Dillman et al. 1998; Hamilton 1999; Dillman 2000; Crawford et al. 2001). A drawback of the invitation letters was that they required potential respondents to type in a web address. A problem with the mass emails was that several were intercepted and deleted by company spam filters and those that got through were often viewed with suspicion by the recipients (our survey unfortunately coincided with widespread publicity about the risk of viruses attached to emails). Only one-third of phone contacts who had agreed to participate actually did so, and the follow-up calls revealed that this was largely due to their having found the website very sluggish. Although website speeds had been tested in the country where the site was hosted, it transpired that the international links were very slow when under heavy load. The slow speeds experienced by users with low-spec equipment or low bandwidth, is a classic source of frustration and low response to web surveys (Solomon 2001).

The hard copy questionnaires were sent out in early 2005 addressed to the Directors of Human Resources of 1,000 UK companies sampled from the FAME database in the same way as had been done for the web questionnaire sample. This produced about 50 completed questionnaires. Telephone calls to a limited number of the non-responding organisations produced further responses but the cost per additional response made it impractical to adopt this method on a large scale.

The low response rates (web $\leq 1\%$, hardcopy $\sim 5\%$) obviously raise the possibility of non-response bias. Although this possibility would clearly undermine the representativeness of any forecasts based on the unweighted data, it presents a useful opportunity to explore the nature of such biases.

## Results

The current article concentrates, not on the results per se, but on their robustness in the face of instrument bias, non-response bias and framing effects and on the risk of myopic interpretation of headline results.

### Respondent characteristics: differences between the two samples

After data cleaning the total sample amounted to 233 questionnaires of which 86 were from the UK (35 web, 51 hard copy), 77 from the Netherlands, 60 from Israel and 10 from Greece. This dataset is not large but is quite adequate for the purposes of this article.

**Table 3** Sample characteristics

| | Web sample | | Hardcopy sample (UK) | Total sample |
|---|---|---|---|---|
| | Non-UK | UK | | |
| Total number of respondents | 147 | 35 | 51 | 233 |
| % Who are part of senior management | 68.5 | 71.4 | 84.3 | 76.2 |
| % Who have used video conferencing in last 2 years | 28.8 | 40.0 | **14.0**** | 27.0 |
| % In groups of ≤50 | **58.7**** | 31.4 | **66.7**** | 56.2 |
| % In groups of >100 | **19.6**** | 40.0 | **21.6*** | 23.2 |
| % In "sales" | 18.0 | 16.5 | 12.9 | 16.7 |
| % In "services to individual clients" | 11.6 | 13.0 | 20.0 | 13.7 |
| % In "services to business customer" | 18.8 | 17.4 | 11.8 | 17.2 |
| % In "retailing/distribution" | 7.7 | 8.7 | 12.9 | 9.0 |
| % With city centre premises | 32.9 | 31.4 | **14.0*** | 28.4 |
| % With isolated/rural premises | 11.0 | 17.1 | 18.0 | 13.3 |
| % With >5% of workforce working from home on a typical day | 25.3 | 20.0 | 21.7 | 23.6 |
| % Where employer encourages working from home | 9.6 | 5.7 | 18.0* | 10.7 |
| % Where employer discourages working from home | 25.3 | 34.3 | 32.0 | 27.9 |
| % With access to video conferencing facilities in the building | 26.7 | 40.0 | **9.8**** | 24.9 |
| % With ≤50 business trips per employee per year | **71.1**** | 88.6 | 84.8 | 76.8 |
| % With >100 business trips per employee per year | 16.2 | 8.6 | 8.7 | 13.5 |
| Average % of workforce working from home on a typical day | 11.7 | 4.9[a] | 10.3[a] | 9.1 |
| Average no. of road-based business trips per employee per year | 50.1 | 35.5[b] | 38.5[b] | 42.7 |

[a] Compares with an independent estimate for UK of 11% for 2005 (though a lower figure might be expected among employees of companies employing at least 10 people) based on an analysis of UK Labour Force Survey data by Ruiz and Walling (2005) assuming that the 11% of the workforce who "usually work from home" do so on 18 days per month and that 4% of the workforce who worked at home on at least 1 day in the survey week do so on 6 days per month

[b] Compares with an independent upper bound estimate for UK of 46% for 2002/3 (though for firms employing at least 10 people a slightly lower figure might be expected)—based on data in tables 3.1 (which reveals 71 business trips per worker per year) and 2.1 (which shows 65% of trips to be by car van or taxi) of DfT (2005)

* Significantly different from the UK web sample (using two tailed *t*-test at 90%)

** Significantly different from the UK web sample (using two tailed *t*-test at 95%)

Some characteristics of the responding population are summarised in Table 3. Given the lack of published data on these characteristics we cannot make any comparison with the wider population but we certainly would not claim to have achieved a representative sample. Table 3 is presented simply as context for our later analyses and to allow comparison of the profile of respondents to the hardcopy version to that of UK respondents to the web version. It appears that, even though the sample frame was the same, the web version is less likely to have been completed by senior management but more likely to have been completed by people with recent experience of using video conferencing. This finding confirms and extends the observation (e.g. by Dillman 2000; Sax et al. 2003;

Arentze et al. 2005) that web surveys are most likely to appeal to people who are computer literate and comfortable with new communication technologies. The organisations represented in the web sample were generally larger, less likely to be involved in the provision of services to individual clients or to be engaged in retailing/distribution, more likely to have access to video conferencing facilities, less likely to encourage home-working, and were likely to generate slightly fewer business-related trips per employee.

Some of the differences are likely to reflect the method of approach (email to general address with request to forward it to Director of Human Resources versus letter directly addressed to that person) and some perhaps to the method of completion (on-line versus hardcopy), but it is salutary to observe how different the two samples appear to be even though they were both drawn from the same database in the same manner and were both intended to be completed by Directors of Human Resources.

Near-comparable published data for the UK is available for the last two rows of Table 3. In each case the hardcopy sample seems closer to the published data than does the web sample. This suggests that, although both methods of recruitment were subject to response bias, the web sample is the more biased of the two.

A further indication of the difference between the two samples is apparent from a comparison, shown in Table 4, of the current and expected future levels of home-working and business trip making reported by respondents in the two samples. Compared to the hardcopy sample, the web sample report much lower current levels of home-working but expect a much more substantial increases by 2010. A similar, but much less marked difference is apparent for business travel, where the web sample report lower current levels and expect a somewhat greater reduction by 2010.

Some consequences of the differences between the two samples are further explored below.

Although the presence of differential response bias has been noted by practitioners of mixed mode surveys (Dillman (2000) quotes a survey undertaken by Biemer (1997) in which the same respondents were surveyed twice using different survey modes and recorded significant differences in their answers to the same question) it is regarded, by Dillman at least, as a price worth paying for the improved overall coverage achievable by using a mixed mode approach. We will discuss this issue in our concluding section.

**Table 4** Expected levels of home-working and business trip making

| | | Respondents' estimate of 2004 level | | Respondents' estimate for the 2010 base scenario | |
|---|---|---|---|---|---|
| | | Web sample (UK & NL) | Hardcopy sample (UK) | Web sample (UK & NL) | Hardcopy sample (UK) |
| % of workforce working from home on a typical day | mean | 5.3 | 10.3 | 12.8 | 9.8 |
| | standard deviation | 9.4 | 25.1 | 16.6 | 15.7 |
| | sample size | 112 | 46 | 105 | 47 |
| Number of business trips per employee per year | mean | 31.5 | 38.5 | 22.5 | 28.8 |
| | standard deviation | 56.9 | 103 | 41.7 | 97.0 |
| | sample size | 111 | 46 | 111 | 46 |

The importance scores

Turning now to the opinions expressed by respondents, and most particularly to their assessment of the importance of factors which had been omitted from the scenario questions. Table 5 suggests that, among the factors listed, respondents think that the most important influences on working from home are speed of communications and security of communications, and that the most important influences on road based business trips are the capability of mobile communications, the presence of high quality video links on most people's desks and, again, the security of communications.

The table also shows consistency between the importance scores given for equivalent factors in the two lists (the two security questions [numbers 1 and 8 in Table 5] had a Pearson's $R$ value of 0.178, the two cost questions [numbers 3 and 10] had a Pearson's $R$

**Table 5** Reported influence of specified factors ($n = 233$)

| | % Reporting the specified factor to have: | | |
|---|---|---|---|
| | "No influence" | "Very minor" or "minor" influence | "Significant" or "very significant" influence |
| *On working from home* | | | |
| 1. Whether security of communications and the internet can be guaranteed | 23.5 | 39.1 | 37.4 |
| 2. Whether the speed of communications links available to a typical home-worker approximates to broadband or is 10 times faster | 20.7 | 24.7 | 44.7 |
| 3. Whether the cost of acquiring and using home office equipment stays at 2005 levels or decreases by 50% | 21.2 | 44.7 | 34.1 |
| 4. Whether journey times stay at 2005 levels or increase by 20% | 24.0 | 45.3 | 30.8 |
| 5. Whether legislation is introduced to remove any liability on employers for accidents or illness attributable to home-working | 25.7 | 51.3 | 22.9 |
| 6. Whether expenditure on the acquisition of home office equipment is allowable against income tax | 20.1 | 45.8 | 34.0 |
| *On business trips* | | | |
| 7. Whether high quality video links are available on most peoples' desks | 19.2 | 43.0 | 37.0 |
| 8. Whether security of communications and the internet can be guaranteed | 20.3 | 44.3 | 35.4 |
| 9. Whether the capability of mobile communications has made the truly mobile office a reality | 18.0 | 44.2 | 37.8 |
| 10. Whether the cost of using communications links stays at 2005 levels or decreases by 50% | 20.3 | 46.5 | 33.1 |
| 11. Whether the cost of travel increases by 10% or decreases by 10% | 23.8 | 57.5 | 18.6 |
| 12. Whether journey times stay at 2005 levels or increase by 20% | 19.9 | 54.9 | 24.6 |

value of 0.693, and the two journey time questions [numbers 4 and 12] had a Pearson's $R$ value of 0.695).

Respondents' statements about the influence of particular factors on the expectations they had reported were compared with their answers to the scenario questions. More home-working was expected in the congested scenario by 38 (69%) of the 55 respondents who said that journey time had a significant or very significant impact on working from home but only by 4 (29%) of the 14 who said that it was of no importance. Similarly, more working from home was expected in the cheap home office equipment scenario by 27 (44%) of the 61 respondents who said that the cost of home office equipment had a significant or very significant impact on working from home, but by none of those who said it was of no importance. Fewer business trips were expected in the reduced communication cost scenario by 35 (67%) of the 52 respondents who said that these costs would have a significant or very significant impact on business travel, but by only 4 (13%) of the 31 who said that it was of no importance. And finally, fewer business trips were expected in the congested future scenario by 29 (71%) of the 41 who said that journey times would have a significant or very significant impact on business travel, but by only 5 (16%) of the 31 who said that it was of no importance. This consistency between the importance scores and the scenario answers provides an internal validation of the latter.

More detailed analysis (not reported here) revealed a number of interesting, and very plausible, differences between the importance scores given by people from different industrial sectors. Overall, there is thus considerable evidence that the importance score questions were being answered consistently. This suggests that respondents were thinking clearly and logically and, by inference, suggests that they were probably also thinking clearly and logically when answering the scenario questions.

This is all very interesting but the key result, in the context of the current article, is that the allocated importance scores indicate that technological factors were perceived as being more important than any of the factors related to cost or journey times. The home-working scenario descriptors (outlined in Table 2) focussed on costs and travel conditions while the business travel scenario descriptors focussed on the price and capability of ICT and on road travel conditions. The spread of high quality video links, and the speed and security of communications were not included in the POET scenario descriptors (except in so far as they can be said to be an aspect of "the capability of ICT"); their exclusion having been justified by the fact that the POET project was not interested in forecasting for futures in which ICT was not fast, secure and ubiquitous. It is of course standard practise to focus the experimental design on the issues of particular interest (e.g. most transport-related SP exercises tend to focus on travel costs and journey durations) but our results highlight the danger of this kind of myopia. The absence of important factors from the scenario variants makes it impossible to estimate their impact on expected levels of home-working or business trip making and leaves any forecasts dangerously exposed to any changes in the omitted variables. Given more time, it would clearly be interesting to repeat the exercise with scenarios dealing with these factors and so see if they do indeed influence stated expectations to the extent that the importance scores would imply.

Models were constructed to explore whether the importance scores quoted by respondents could be explained by any of the attributes of the respondents or of their organisations. Although none of these models achieved $R$ square coefficients in excess of 0.21 they did suggest that different factors were important to different types of respondent (e.g. that respondents who use video conferencing tended to give high importance scores to the capability, security, cost and penetration of ICT equipment but to give low importance scores to the cost of duration of journeys).

A comparison (not shown here) of regression models built with and without the respondents' importance scores confirmed that the expected levels of home-working and of business trip making could be better explained by including the importance scores (the $R^2$ values increased by 0.07 and 0.15, respectively when the importance scores were allowed to be included as explanatory variables). Models of home-working were improved by including the respondent's assessment of the importance of speed of communications, cost of communications, liability legislation and cost of home office equipment. Models of business trips were improved by including the respondent's assessment of the importance of having high speed video links on most people's desks.

## The influence of preview texts

As will be recalled, an innovative feature of our questionnaire was the inclusion of texts designed to test the robustness of respondents' opinions in the face of outside influences—the premise being that robust opinions will be relatively unaffected by outside opinions while more superficial opinions will be more easily swayed (Biek et al. 1996; Visser et al. 2006). Each respondent was randomly assigned one block of text which was either generally positive or generally negative about working from home and another which was either generally positive or generally negative about ICT as a substitute for business trips (see Table 6). Respondents were then asked to assess the degree of balance in the texts they had received.

**Table 6** Issues addressed in the briefing texts

| | Briefing on tele-working | Briefing on ICT as substitute for business travel |
|---|---|---|
| Positive | Employee appreciation of the opportunity to work around their homecare responsibilities and the reduced amount of traveling;<br>Employer benefit from reduced costs of premises, increased employee satisfaction and access to a larger, computer-literate, labour market;<br>The opportunities provided by the latest generation of ICT for conferencing and close monitoring of off-site workers | Rapid improvement in the quality of desk-top video conferencing and the ability to access such services without pre-booking;<br>Availability of ancillary services to provide instant access to documents and meeting transcripts;<br>The upcoming generation's familiarity with ICT;<br>The modern sharp image projected by companies which use ICT in place of face-to-face meetings;<br>Ease with which additional staff can be included in "meetings";<br>Increases in efficiency due to more focused meetings and reduced time and money spent traveling. |
| Negative | Employees getting distracted from their work, missing the companionship of the workplace and resenting the cost of their home office;<br>Employer's liability for injuries suffered by employees working from home;<br>Reduced corporate spirit and company loyalty;<br>Security considerations. | Customers' appreciation of the effort taken to set-up and attend face-to-face meetings;<br>The importance of human contact;<br>The fact that developments in route guidance and mobile-office technology have made it possible to deploy field staff much more efficiently;<br>The risk of equipment failure and security breaches. |

**Table 7** Assessment of balance in the briefing texts ($n = 233$)

| Respondent's assessment of the text | Percentage of the group who received the text which was… | | | |
|---|---|---|---|---|
| | … Positive about working from home | … Negative about working from home | … Positive about ICT as substitute for business trips | … Negative about ICT as substitute for business trips |
| "Too enthusiastic" | 49.8 | 4.4 | 49.4 | 18.9 |
| "Well balanced" | 45.3 | 48.6 | 43.2 | 66.4 |
| "Too negative" | 4.9 | 47.0 | 7.4 | 14.7 |
| Total number in the group | 128 | 105 | 132 | 101 |

Table 7 summarises the respondents' assessments of the texts and shows that, whereas scarcely 10% of respondents held more extreme views than those expressed in the home-working texts (with similar proportions thinking the negative text too enthusiastic and thinking the positive text too negative), about 26% of respondents held more extreme views than those expressed in the ICT-as-a-substitute texts (with a particularly high proportion thinking that the negative text was too enthusiastic). This suggests that, while our home-working texts seem to have reflected the spectrum of our respondents' opinions, our ICT-as-a-substitute texts were probably positioned towards the "too enthusiastic" end of that spectrum.

Models (not shown here) were constructed in an attempt to explain the respondents' assessments of the briefing texts. The models did not provide a high level of explanation ($R^2$ less than 0.23) but suggested that the text which had been positive about working from home was thought to be too enthusiastic by respondents who themselves worked from home and by those whose employers discourage it, but was thought to be too negative by respondents who were part of the senior management team or had numerous colleagues working from home. The text which had been negative about using ICT to replace business trips was thought to be too negative by respondents who were part of the senior management team or whose companies were involved in supplying financial services. These are interesting findings.

The most important result in Table 7 is contained in the "totals" row, which shows that we have received fewer responses from people who received the negative texts. Given that an equal number of positive and negative texts were assigned this suggests that receipt of a negative text (particularly the text about ICT as a substitute for business trips) reduced the probability of the respondent completing and returning the questionnaire. This result was not welcomed, because it upset the symmetry of the experiment, but is important because it draws attention to a particular form of response bias whose implications are rarely considered by survey designers (see discussion below).

Table 8 tabulates some other effects of the briefing texts. It appears that the texts had an influence on some, but not all, of the importance scores. For example, respondents who received the positive were more likely to report that journey times are an important factor in the amount of home-working, and that the presence of high quality video on most people's desks is an important factor in the substitution of business trips (note that both of these issues were alluded to in the positive texts but not in the negative texts).

**Table 8** Influences of briefing texts on respondents' attitudes and expectations

| | Among respondents who received text which was: | |
|---|---|---|
| | Positive | Negative |
| *Percentage of those who thought the following were very important in influencing the numbers who would work from home in 2010 (n for positive = 128, n for negative 105)* | | |
| Security of communications and internet guaranteed | 4.4 | 4.4 |
| Communication links are ten times faster | 5.9 | 5.2 |
| Costs of home office equipment stay the same or reduce by 50% | 4.4 | 4.4 |
| Journey times remain same or increase by 20% | 3.7** | 1.5** |
| Legislation removes employers' liability | 3.0 | 3.7 |
| *Percentage of those who thought the following were very important in influencing the numbers of business trips in 2010 (n for positive = 132, n for negative 101)* | | |
| High quality video links are on most people's desks | 4.3* | 1.9* |
| Whether security of communications and the internet can be guaranteed | 6.3 | 4.3 |
| Does capability of mobile communications make the mobile office a reality | 6.5 | 4.3 |
| Costs of using communications stays the same or decreases by 50% | 2.5 | 3.0 |
| Whether cost of travel increases of decreases by 10% | 1.3 | 1.2 |
| Journey times stay the same or increase by 20% | 3.3 | 3.0 |
| *Effect on reported levels of home-working and business travel*[a] | | |
| % of workforce thought to be working from home on an average day in 2004 (*n* for positive = 100, *n* for negative 91) | 13.2* | 16.5* |
| Estimated number of business-related trips per employee in 2004 (*n* for positive = 112, *n* for negative 89) | 36.3** | 54.2** |
| *Effect on expected levels of home-working (n = 191) and business travel (n = 201)*[b] | | |
| % of workforce thought likely to be working from home on an average day in 2010 (*n* for positive = 100, *n* for negative 91) | 19.7 | 19.1 |
| Expected number of business-related trips per employee in 2010 (*n* for positive = 112, *n* for negative 89) | 32.4 | 37.3 |

[a] Excludes data from respondents who did not answer the scenario questions

[b] Based on response to the base ("no change") scenario for 2010

* Two tailed *t*-test reveals difference significant at 10%

** Two tailed *t*-test reveals significant difference at 5%

Results in Table 8 suggest that the texts have influenced the average reported levels of home-working and business travel. These effects (which are significant at 10% and 5%, respectively) were totally unexpected—particularly since the questions on current levels of these activities preceded the briefing texts. A possible explanation is that the negative texts' depressive influence on response rates (referred to earlier) may have been stronger for respondents who thought current levels of home-working and business travel were low—thus increasing the average reported levels. Whatever the reason, the effects of the briefing texts were more pervasive than we had anticipated. We discuss the implications of this in the concluding section.

The final two rows of Table 8 show the effect of the texts on expected 2010 levels of home-working and business-related travel respectively. The effects are in the intuitive

direction (with the positive texts apparently leading people to expect more use of ICT and less travel) but are not statistically significant. Regression models (reported in Bonsall and Shires 2006) which allowed for interaction effects showed that, although the texts had no significant effect on estimates of future home-working, they did have a significant moderating effect on the predicted influence of scenario variables on future business trip making.

### Differences in models based on data from the two sources

It was noted earlier that the web-based and hard copy samples reported different current levels of home-working and of business travel and that they had different future expectations for these activities. Some of this difference may be simply explained by differences in the sample make up (e.g. respondents in the web sample tended to be less senior and to come from larger organisations). However, models of the factors affecting current and expected future levels of home-working and of business trip making suggest that this is not a complete explanation.

Our comparison of models built on data from the web sample with models built on data from the hard-copy sample (reported in Bonsall and Shires 2006) revealed that, although both data sources reveal factors which have the same direction of influence,[1] the magnitude of this influence is different in the two data sets and that the two data sources differ in respect of all the other statistically significant influences and, in some instances, even differ in respect of the direction of that influence.

Among other differences, it is interesting to note that the web sample's estimates of levels of home-working and business-related travel under the different scenarios were more influenced by technological factors such as the speed and security of communications. Similarly, it is interesting to note that the influence of the briefing texts on estimates of business-related travel under the different scenarios was significant within data from the web sample but was not apparent in data derived from the hard-copy sample.

Overall, the models based on data from the hardcopy questionnaire provided a better explanation of reported current (and future expected) levels of home-working and business travel than did those based on data from the web questionnaire.

## Discussion and conclusions

This article has used an unusual data set to explore a number of issues of general interest to survey designers and users of questionnaire data. The unique design of the POET questionnaire, and the fact that it was administered in two versions, has made it possible to explore a number of issues which are rarely considered in such detail. Although the achieved sample would make it unsafe to claim that the specific results are representative, we believe that several of the broader findings are relevant to the collection and analysis of data in other contexts.

[1] In the models of future levels of home-working, both data sets reveal the positive influence of the current level of home-working, increases in journey times, decreases in home office costs, and the negative influence of having city centre or town centre premises. In the models of future business travel, both datasets reveal the positive influence of the current business mileage driven and the negative influence of the price and capability of ICT, the seniority of the respondent, and involvement in sales.

Survey mode effects

Experience with the web version of the questionnaire highlighted several problem

- Recruitment to a web survey is not always straightforward (mass emailing can, if not personalised, be regarded as spam and produces a very low response; recruitment via mail is not effective—probably because it requires respondents to be motivated to make the effort to visit the website; and although telephone recruitment can be effective, it is very expensive). The use of invitations and hotlinks posted on other websites may be effective—and can be arranged at some cost through specialist web-marketing organisations, but is clearly not appropriate unless the desired sample is likely to be visiting those sites.
- It is clear that, because of widespread fear of computer viruses, many potential respondents are reluctant to visit a website whose credentials they have to take on trust.
- Software/hardware problems can contribute to a reduced response rate if they make access to the questionnaire slow or unreliable (complex questionnaire software may result in a slow response time and problems with the reliability of website links may not be predicted or easily resolved).
- Because of these problems of recruitment, the cost savings which may be expected to be associated with web surveys may be illusory (particularly when the costs of preparing the web questionnaire are taken into account).

The use of a hardcopy version of the questionnaire alongside the web version allowed us to make some interesting comparisons. It is clear that, *even when the same sampling frame was used*, the respondents to the two versions of the questionnaire differed in several respects. Compared to the respondents to the hardcopy version, respondents to the web version tended to be more junior, to come from larger organisations, to have more recent experience of video conferencing. These differences seem to have led them to put more weight on technology-related factors when estimating future levels of home-working and business travel under different scenarios (perhaps, being more technologically oriented, the web respondents were more likely to see effective technology as an important pre-requisite for home-working and for substitution of business trips by ICT), and to have less robust prior opinions about future levels of business travel (perhaps, being more junior, they had less experience of the issues influencing the volume of business travel). Whether the differences between the two samples mean that the web sample is more, or less, useful is of course open to discussion but it is important to recognise that the samples are different. In our study the respondents to the web version expected higher levels of home-working, and of substitution of business travel by ICT, than did the respondents to the hardcopy version. Naive forecasts based on these expectations would clearly differ depending on the method of data collection employed.

Some authorities (e.g. Dillman 2000) advocate the use of a mixed mode survey strategy in order to maximise coverage of the population. Dillman acknowledges that responses to different surveys modes may differ but suggests that little can be done to overcome the problem and that it is a price worth paying in pursuit of improved coverage. We cautiously endorse that opinion but emphasise the need for careful analysis of the different results achieved from the different sub-populations.

Some authorities (e.g. Dillman 2000; Kaplowitz et al. 2004) have suggested that questionnaires administered via different media may yield different answers because the respondents will experience the questions in different ways because of framing effects and subtle differences in presentation (e.g. the fact that, with a hardcopy version, respondents

can look ahead or take questions out of order whereas, with a web version, this can be—and in our case was—prevented). Our work leads us to conclude that the differential response issue (which is a form of selection bias) may actually be much more important than the framing issue.

### The use of supplementary questions seeking respondents' opinion on the importance of "extra" factors

Detailed analysis of the questions which sought respondents' views on the importance of specified factors on their expectations of future levels of home-working and business travel indicated that technological factors were generally thought to be more important than factors related to costs or journey times. Given that Stated Preference (SP) and Stated Expectation (SE) exercises traditionally focus on costs and travel conditions, this result highlights the value of additional questions of this kind. Such questions should ideally be used in a preliminary study to help establish the variables to be included in the SP/SE experiment. However, as we have seen, they can also play a useful role even after the experimental design has been fixed; they can provide an internal validity check on the SP results and can serve as a useful reminder of the limits to the domain in which the project forecasts are based. That is to say, they remind the analyst not to forget the wider focus.

In our case the inclusion of the importance scores among the explanatory variables of our models of respondents' expectations resulted in a significant improvement in the explanatory power of the models. This highlights the extent to which opinions and assumptions influence public expectations about home-working and business travel. This result, together with our conclusion (see below) that opinions on home-working and business travel are quite labile, points to the extent to which publicity and marketing might influence public expectations in this sector.

### The use of briefing texts to test robustness of opinions

The most innovative feature of our questionnaire was the use of briefing texts to test the robustness of respondents' opinions in the face of outside influences. This feature has allowed us to draw three important conclusions

- The tenor of the briefing texts appears to have influenced the response rates; inclusion of a negative text, particularly one relating to ICT as a substitute for business travel, seems to have reduced the likelihood of the questionnaire being completed (alternatively, the positive text may have increased the response rates).
- The tenor of the briefing texts appears to have had some influence on the average reported levels of current business travel and, to a lesser extent, of home-working; demonstrating that an unbalanced briefing could distort the data in unexpected ways.
- The tenor of the briefing texts appears to have had some influence on respondents' stated opinions and on their expectations, under future scenarios, of levels of business travel and of home-working; although many of the effects are not statistically significant it is clear that the expressed opinions and expectations are labile and that some are more labile than others.

Our finding that the tenor of briefing texts can influence response rates might tempt survey organisers to use positive briefings as a way of increasing their response rates. There are two reasons for discouraging such a strategy. Firstly, because, given the existence of widespread lability, a positive message is likely to influence, and thus bias, the

opinions expressed by respondents; and secondly because, since individuals are less likely to respond to questionnaires which are associated with values which they do not themselves espouse (e.g. Eagly and Chaiken 1993), the use of positive messages is likely to decrease the likelihood of responses from people who hold negative opinions—thus biasing the data in potentially complicated ways. Taken together, these two tendencies could seriously distort the spectrum of behaviour and opinion represented in the questionnaire results.

Lability of opinions and expectations is a problem for forecasters and modellers. Our concern is that such lability is usually invisible to the analyst but that ignoring it does not make it go away; had we not tested the effect of contrasting briefings we would not have been aware of the lability or of the fact that it was more marked in some parts of the data set than in others. Part of the unexplained variance in the data might have been associated with lability but we would have had no way of knowing how much. If we had given all our respondents the same briefing material, or none at all, the variance in the data would have been reduced and better fitting models might have been produced. But, if lability exists in real life, this improvement in the models would have been fragile and illusory.

An observed change in opinions recorded at different points in time may be indicative of lability but, not only is it difficult to know whether an observed change is likely to persist, but it may not be practical to delay the analysis until results from a series of surveys is available. The possibility that lability could be detected via a single survey is obviously attractive and the use of contrasting briefing texts to test the robustness of respondents' prior opinions offers such a possibility. If contrasting briefings are found to have no effect on the data then the analyst could be quite confident that lability is not a serious issue, but if there is an effect then this should be reported and appropriate actions—such as sensitivity analyses—taken to explore its implications.

The impact of contrasting briefings will, of course, depend on how far apart the opinions in the two texts lie. Use of a standard degree of separation (e.g. a positive text thought too enthusiastic by 75% of a pilot sample and a negative text thought too pessimistic by 75% of that sample) might offer one way forward but it may, in practice, be difficult to afford a pilot sample which is large enough to overcome any potential response bias and establish these statistics.

Another practical issue may be that the sponsors of the survey may feel uncomfortable with the idea of being associated with the "negative" version of the briefing (imagine a local authority wishing to gauge support for a road pricing scheme—would they be happy to sponsor a survey which drew attention to the negative aspects of the scheme?). This brings us to an important issue—the importance of preserving balance in the overall briefing. Standard guides to best practice in survey design, supported by research (see for example Dillman 2000; Sommer and Sommer 2001; Groves et al. 1992), emphasise the importance of a balanced presentation. We do not dissent from this advice but suggest that it should apply to the whole population sample rather than, necessarily, to each individual respondent.

In fact we doubt that balanced presentations are always achieved in practice. A lack of balance may sometimes be attributable to lack of care; sometimes it may be an unintended consequence of an attempt to encourage a high response to the survey (e.g. in order to achieve a good response to a survey on cycling and walking, the briefing might indicate the importance of such data in estimating future levels of obesity—thus drawing attention to a positive aspect of walking and cycling); and sometimes it may be because the survey sponsor has a biased idea of what constitutes a balanced presentation.

More fundamentally, against a background of varied and potentially labile opinions, it is difficult to be certain what is really meant by a balanced presentation. If we define balance as having been achieved if the number of people regarding the text(s) as too positive is similar to that regarding it/them as too negative, then questions similar to our questions B2/C2 could be used during initial piloting to establish whether this is the case. However, if public opinion is labile, the result of such a test will simply reflect recent events and publicity rather than provide a measure of "true" balance. The use of contrasting briefings would, again, offer a means of ascertaining whether opinions were labile and thus whether any estimate of balance could be made.

We suggest that all these issues would benefit from further research.

**Acknowledgments** We wish to thank colleagues in the POET consortium, most notably Gerard de Jong, for encouraging comments during development of the questionnaire and during our model development. Thanks are also due to Dimitris Karalis (UoA) and Jason Bennet (SurveyEngine) for their operationalisation of the web version of the questionnaire and to Nusrat Walid for assistance in UK survey administration. We would like to thank the anonymous referees whose comments on an earlier version of this article have been most helpful. Finally we wish to thank the European Union and The Institute for Transport Studies for their financial support.

## References

Arentze, T.A., Bos, I., Molin, E.J.E., Timmermans, H.J.P.: Internet-based travel surveys: selected evidence on response rates, sampling bias and reliability. Transportmetrica **1**(3), 193–207 (2005)

Biek, M., Wood, W., Chaiken, S.: Working knowledge, cognitive processing, and attitudes: on the determinants of bias. Pers. Soc. Psychol. Bull. **22**(6), 547–556 (1996)

Biemer, P.P.: Health Insurance Finance Agency Evaluation. Unpublished data. Research Triangle Institute, Research Triangle, NC (1997)

Bonsall, P.W., Shires, J.: Employer expectations for commuting and business-related travel in an environment rich in information and communication technologies. Transportation Research Record 1977, pp. 268–276 (2006)

Crawford, S.D., Couper, M.P., Lamias, M.J.: Web surveys; perception of burden. Soc. Sci. Comput. Rev. **19**(2), 146–162 (2001)

De Jong, G., Algers, S., Papola, A., Burg, R.: Impact of e-Economy on traffic and traffic-related indicators in urban areas. Transportation Research Record 1977, pp. 286–291 (2006)

DfT: "Focus on Personal Travel", Analysis of Data from the UK National Travel Survey, 2005 Edition. Department for Transport, London (2005)

Dillman, D.A.: Mail and Internet Surveys: The Tailored Design Method, 2nd edn. John Wiley and Sons Inc., New York (2000)

Dillman, D.A., Tortora, R.D., Conradt, J., Bowker, D.: Influence of plain versus fancy design on response rates for web surveys. Unpublished paper presented at annual meeting of the American Statistical Association, Dallas, TX 1998

Eagly, A.H., Chaiken, S.: The Psychology of Attitudes. Thompson Wadsworth, Belmont, CA (1993)

Groves, R.M., Caldire, R.B., Couper, M.P.: Understanding the decision to participate in a survey. Public Opin. Q. **56**, 475–495 (1992)

Hamilton, J.C.: The ethics of conducting social-science research on the Internet. Chron. High. Educ. **46**, B6–7 (1999)

Kaplowitz, M.D., Hadlock, T.D., Levine, R.: A comparison of web and mail survey response rates. Public Opin. Q. **68**(1), 94–101 (2004)

POET Consortium. Results from virtual case studies. POET Project Deliverable 9. Available for download from the POET website: http://www.poet-eu.org/documents.asp (2005a)

POET Consortium. Results of six case studies on urban transport. POET Project Deliverable 10. Available for download from the POET website: http://www.poet-eu.org/documents.asp (2005b)

Ruiz, Y., Walling, A.: Home-based working using communication technologies. Labour Mark. Trends **113**(10), 417–426, Office for National Statistics, UK (2005)

Sax, L.J., Gilmartin, S.K., Bryant, A.N.: Assessing response rates and non-response bias in web and paper surveys. Res. High. Educ. **44**(4), 409–432 (2003)

Schaefer, D., Dillman, D.A.: Development of a standard e-mail methodology: Results of an experiment. Public Opin. Q. **62**, 378–397 (1998)
Smith, C.B.: Casting the net: surveying an Internet population. J. Comput. Mediat. Commun. **3**, 77–84 (1997)
Solomon, D.J.: Conducting web-based surveys. Pract. Assess. Res. Eval. **7**(19) (2001)
Sommer, B., Sommer, R.: A Practical Guide to Behavioural Research: Tools and Techniques, 5th edn. Oxford University Press, New York (2001)
Tversky, A., Kahneman, D.: The framing of decisions and psychology of choice. Science **211**, 453–458 (1981)
Visser, P.S., Bizer, G.Y., Krosnick, J.A.: Exploring the latent structure of strength-related attitude attributes. Adv. Exp. Soc. Psychol. **38**, 1–67 (2006)

## Author Biographies

**Peter Bonsall** Professor of Transport Planning at the Institute for Transport Studies, University of Leeds. His research interests include: use of innovative data sources, microsimulation, multi-criteria appraisal of policy interventions, travellers' perception of modal attributes, their ability to cope with uncertainty and complexity and their response to new information and charges.

**Jeremy Shires** Senior Research Fellow at the Institute for Transport Studies, University of Leeds. His research interests include behavioural modelling, the impact of "soft factors" on travel, stated preference design and public transport demand modelling.