

## 6. Workshop report: good data is key to the development of good models: so how is innovation in data collection keeping apace?

**Rob Sheldon, Martin Dix, Terry Flynn and Paul Metcalfe**

---

### INTRODUCTION

The market research/survey industry has, for many years, really been seen as the 'poor relation' of other disciplines such as modelling and economics.

Yet there are many ideas that have been and continue to be developed from such disciplines, and it is critical to ensure that the data collection side of things can keep pace with them. Equally it is clear that market research/survey practitioners also have ideas that it would be useful to cross fertilise with other disciplines, to produce more cost efficiency, innovation and quality enhancement.

This was the basis for recommending the workshop for the conference agenda. It was designed to provide a useful complement to the complex and innovative modelling issues that provided the main fare for the conference. The workshop was designed to run for an entire morning session. The following programme was followed:

- Introduction: Rob Sheldon (Accent, Chair)
- How qualitative research is putting itself at the heart of the process: Martin Dix (MVA Consultancy)
- Design and cost efficiency developments in quantitative data collection – keeping the relationship dynamic: Rob Sheldon
- Is there more to come through the demands of increased modelling sophistication? Terry Flynn (CenSoc)

Dr Paul Metcalfe then set out a number of issues as a discussant to pump prime the debate. There were around twenty participants to the workshop

which generated a lively discussion. Each of the presentations is discussed in turn. This is followed by a section summarising the key themes that emerged from the debate.

## 1. HOW QUALITATIVE RESEARCH IS PUTTING ITSELF AT THE HEART OF THE PROCESS

### 1.1 Background

The first presentation began by noting that recent developments in both qualitative and quantitative survey methods are improving the efficiency of research through cheaper, faster information gathering. Researchers working in most fields today are, however, increasingly facing new challenges, including consumer resistance to surveys, reflected in falling response rates; increasingly tactical attitudes to answering survey questions and consumer disinterest in 'standard' forms of survey. These factors are driving the need for researchers to take more engaging and innovative approaches to survey design.

At the same time research buyers, including many UK government bodies, increasingly want more than just the outputs from classic 'what-if' predictive models. There is increased demand for clear practical guidance on what actions they should take to influence consumer behaviour, effectively and efficiently. For example: delivering clear strategy on how to achieve switch in travel behaviour from car to more use of sustainable options, including 'active travel' modes. This has fuelled interest in research approaches that can combine the strengths of quantitative and qualitative methods, to deliver reliable forecasts and insightful understanding of how consumers actually choose, and how to influence their decision processes.

In the field of choice modelling, qualitative research has often had an important role in guiding quantitative and particularly Stated Preference (SP) survey design, particularly through cognitive testing and diagnostic questions. It has thereby contributed to improvement in the quality of data, model outputs and depth of insight we are getting out of choice model based research.

However in the author's experience some modellers do not really fully appreciate what qualitative research is and how it can help them. The presentation therefore attempted to address this lack of understanding by exploring the questions: why should choice modellers need qualitative research? How can qualitative methods, especially newer approaches and emerging online techniques, help meet today's challenges and enhance the survey designer's toolkit?

‘Qualitative’ methods were taken to include:

- exploratory and elicitation methods to assist good design of SP and quantitative survey approaches, the classic methods being focus groups and depth interviews
- cognitive testing of survey instruments during their design and development stages
- diagnostic questions, during or following questionnaire completions, to check the validity of responses and gain additional insight
- deliberative methods, which iteratively explore changes in consumer attitudes or behaviour intentions in response to information about the policy, its potential impacts, and how other consumers react to or feel about the same information, to help understand key drivers of attitude/behaviour change
- emerging use of online discussion forums, ethnographic blogs and social network data

Five examples of ‘good practice’ were next outlined, each illustrating different roles for qualitative research methods as part of the overall research project and their benefits for the choice modeller.

## 1.2 ‘Good practice’ case study examples

### **Example 1: Helping to understand decision processes so that models can reflect them**

There has long been recognition that Random Utility Model (RUM) representations of decisions may not reflect real life decision making in some consumer areas or by every consumer. Arguably this would not matter if model predictions were acceptably accurate, but we seldom know whether or not they are. Accuracy of predictions might be improved if we could identify any choice processes, or types of chooser, that do not conform with RUM formulation in the use of qualitative techniques.

A paper presented to the Inaugural ICM conference by Hensher (2009) highlighted the importance of recognising that consumers often apply heuristics and may not therefore act as perfectly informed utility maximisers. Heuristics include strategies for simplifying choices, including non-compensatory option selection/elimination. The question arising is: how can these be identified?

Qualitative methods have some advantages in ‘getting into the respondent’s head’ to understand their choice process, in the context being investigated, while reducing risk that the understanding gained is based on prior assumptions or preconceived theories.

Bonsall et al. (2006 and 2007) used qualitative research methods to investigate consumers' likely responses to potentially complex 'road pricing' tariffs and other (real) tariffs such as mobile phone charging. The work suggested that 'a significant proportion of consumers "disengage" if they perceive cost structures to be too complex; this disengagement sometimes leads them to avoid exposure to that cost but sometimes leads them simply to pay up regardless'. This finding had important policy implications. It suggested that price signals intended to achieve reduced car use at more congested times and busier parts of the network would achieve smaller behaviour shifts (though greater revenue generation) than conventional models would predict. The research also identified distinct cognitive types ('determined', 'cautious' and 'trusting') that were linked (loosely) with age, gender and education. The study conclusions included recommendations for the modelling team to recalibrate and consider the potential for segmentation to recognise cognitive type.

**Examples 2 and 3: Informing segmentations – by cognitive/decision style, attitude profile, and psychographics, not just demographics**

Qualitative research grounded in attitude-behaviour theory can help to identify important behaviourally distinctive consumer segments. A good example is Jillian Anable's classification (Anable 2005) of people resident in any given area according to their 'propensity to shift from car'. This is a useful quantitative segmentation developed from a qualitatively derived typology. It has enabled the numbers and locations of residents who fall into each segment (eg. diehard drivers, reluctant riders, car aspirers, car sceptics, car complacent, aspiring environmentalists and malcontented motorists) to be identified, thus helping to guide targeting of interventions. It has been influential in the UK for planning 'smarter travel' campaigns.

Crockett and Whelan's work (Crockett et al. 2008) on attitudes to public transport modes in South Yorkshire, UK, to help plan for maximum take-up of a Superbus, provides another example of the use of qualitative research to help develop psychographic segmentations that differentiate people with varying propensity to change their behaviour. A battery of psychographic profiling questions, developed through qualitative research, captured individual beliefs concerning such issues as: climate change, importance of transport for climate change, whether the respondent's decisions impact on others and strength of concern about traffic and congestion levels. Distinctive segments emerged and since membership of particular segments could be linked to demographic and socio-economic characteristics, it was possible to develop thematic mapping using available census data to guide the efficient targeting of public communications.

**Examples 4 and 5: Informing and guiding key aspects of SP survey design**

Qualitative research can have key roles during the design and development stages of SP surveys in: selecting and defining attributes and levels; communicating (hypothetical) alternative options/packages meaningfully; identifying 'reference alternatives' and selecting the most suitable method for capturing preferences (whether ranking, rating or pairwise comparison).

Two 'case study' examples illustrate the role of qualitative research in improving the accuracy and value of SP study outputs.

A study in 2010 by MVA Consultancy for the UK Department for Transport (MVA Consultancy 2009) aimed at quantifying the value rail passengers would attach to reductions in overcrowding. A key challenge was to depict, realistically and objectively, a range of rail crowding levels including some that most respondents would never have experienced, enabling them to compare their anticipated journey experiences and consider willingness to pay more for preferred situations. The work began with focus groups, leading to cognitive testing, computer simulations and pilot surveys.

The focus groups generated a wide range of detailed conclusions (for example, simplified visual charts were more unambiguously interpreted than photographs, which introduced distortions and emotive distractions; participants liked to be able to state 'strongly prefer' as well as 'prefer', and fatigue effects meant that no more than four or five scenarios should be presented). The success of study outputs, in estimating robust choice models generating plausible values which were easy to apply within existing modelling frameworks, was in part due to meticulous crafting of SP experimental design using qualitative research to make it relevant to the respondent and at the same time provide an objective measure of crowding.

Another study by MVA Consultancy for the UK Department for Transport (Le Masurier et al., 2007) aimed to obtain UK residents' valuations of reductions in aircraft noise. Again, a key challenge was in realistically communicating the experience of possible future aviation noise patterns. Scenarios had to cover variations in: total number of fly-over events in the period; mix of aircraft types (some louder, some quieter, with different noise qualities) and different periods of day and night. Focus groups and depth interviews helped guide the development of high quality audio-visual stimulus material. This included actual noise recordings taken outside each respondent's home, played back inside their home at calibrated level. Noise profiles for each aircraft type were referenced to the aircraft names with photos and line drawings. Cognitive testing was used to assess the most suitable format of SP option presentations and methods of capturing preferences. This found that comparing three configurations,

each indicating aircraft numbers, aircraft type and 'annual grant' (which of these three is best, which worst), was most respondent friendly and metrically powerful.

Diagnostic questions were also developed for this study, and asked during and after each main survey interview, to assess the validity of responses. While these largely confirmed that respondents had understood the tasks, followed instructions and stated their preferences validly and accurately, they did however also point to issues that highlighted limitations of the SP approach in respect of the study objectives.

Among them, although respondents were asked 'which would be the best situation for you and your household', few could answer reliably for 'their household'. This was an important limitation as household level valuations (through taxation or grant mechanisms) were required. Of greater concern, many respondents indicated that in the longer term they would probably 'get used to' the noise changes they had considered and that the feeling of nuisance from unfamiliar noise sources would reduce in time.

The suggestion that habituation will occur implies that in reality, over time, consumers' tolerance thresholds will rise, so the value of improvements in (in these examples) aircraft noise and rail crowding levels will be less than that identified by the research. This is evidenced, in respect of tolerance for rail crowding, by the much steeper rail crowding curves for areas of the UK outside of London and the Southeast, where crowding is less prevalent.

### **1.3 Implications: limitations of conventional SP survey measures**

These findings from diagnostic questions point in turn to some wider limitations of most conventional SP survey methods, to which researchers should be alert.

First, respondents are usually questioned in isolation, which can mean that influences of other people on their real-life choices (importantly the influence of family and peers) may not be fully captured.

Second, it is important that respondents' preferences are based on their having accurately informed awareness of each alternative option, and considerable effort is properly devoted to achieving this (as illustrated in the 'good practice' examples cited). However to anticipate the real life market response to product/service changes it is important to take additional account of the dynamics of response to change, recognising that (for example), in the marketplace, initially limited awareness will be followed by diffusion of knowledge about the product/policy/service options, and the experiences of 'early adopters' will tend to shape the attitudes/intentions of others, by influencing norms and beliefs.

In addition, as already noted, any tendency for consumers to habituate/adapt to initially unfamiliar and initially strongly valued attributes/stimuli (such as changes in noise or crowding) is not easily captured by conventional SP surveys which, by their nature, focus on measuring instantaneous response to step changes, attempting to ensure that survey respondents have perfect information on each option.

#### **1.4 Emerging developments in qualitative research**

The final part of the presentation therefore addressed the questions: can qualitative research techniques, particularly new and emerging methods, help to address issues and limitations of classic SP based choice models and enrich their outputs? What extra dimensions can they bring to understanding/modelling choices?

A brief account was given of three relatively new, innovative and emerging approaches.

##### **Deliberative workshops**

Deliberative workshops became popular among government and other UK research clients in the last decade as a way of exploring public response to possible policy interventions. The approach is often aimed at identifying the key policy levers and communications messages that might effectively influence the desired change in consumer attitudes and/or behaviour.

A common format is to organise one or more one-day sessions with a specifically recruited panel or cross-section of consumers. Sessions explore, through a step-by-step approach, changes in initial attitudes and behavioural intentions (such as attitudes to using the car) in response to presentations of information about the background to the policy (such as evidence on climate change), its potential impacts (often through presentations of research findings and/or expert projections), and experience of how other consumers are reacting and feeling (through discussion and often mini-polls at intervals during each session).

Deliberative workshops can be considered a useful complement to conventional SP choice modelling because they can generate insights into aspects of consumer choice that the latter are seldom designed to explore, or able to investigate. In particular:

- identifying which types of consumer are more likely to be early vs. late adopters (as these may often be very different segments behaviourally and in their characteristics)

- exploring learning curves and tipping points (so that these might be captured in a demand forecast model)
- understanding how the supply of information on options affects choices (to help identify key segments for targeting interventions and key messages for public communications)
- seeking to identify key drivers of change in attitudes or behavioural intention that are actionable by policy makers
- attempting to disentangle the influences of: respondents' characteristics, their past experience, and their responses to information presented and other respondents' responses.

These methods may have particular value at the post-modelling stage, to test and challenge predictions or assumptions in demand models and to gain additional insights which may assist policy implementation.

### **Online discussion groups**

The surge in use and availability of online panels for quantitative surveys, a subject addressed by the second presentation in this workshop, has also generated increasing interest in online discussion groups, though as yet there has been little take-up, at least by transport researchers.

These may be one-off or repeated, recruitment can be targeted and the method is relatively cheap and efficient. Since (unlike conventional focus groups) responses to questions can be obtained from all in the group simultaneously, it is possible to capture detailed information in as little as half an hour. However online groups may also be structured to follow a stepwise approach, over several days or weeks, similar to the deliberative workshop method.

The main limitation, apart from issues of the representativeness of online panel recruitment, is that since contributions are written they are more variable in quality than those from a focus group, where both verbal and non-verbal responses can be observed. Advantages of the online method include the facility for incorporating good quality audio-visuals, which can itself increase interest and engagement, and the instant generation of transcripts for analysis.

### **Using social network data and 'ethnographic blogs'**

Researchers in some consumer research fields now routinely monitor chatter on social network sites to track brand perceptions, and sites like Facebook can be used to host discussion forums.

Additionally participants in otherwise conventional research (e.g. classic focus groups) may be asked to blog their experiences (e.g. for periods before and after the live session) enabling researchers to build profiles and

track the impacts on ongoing attitudes/behaviour of information messages presented in the focus group.

As yet however these approaches have had little impact on research practices in the choice modelling community and their potential remains to be seen.

### **1.5 Closing comments**

Concluding the first presentation, it was noted that qualitative research has long had a key role in informing the design of quantitative and SP surveys and enriching their outputs, yet some choice modellers have not fully appreciated what qualitative research is and how it can help them. It was suggested that in the workshop discussion session it would be useful to debate: what are the key benefits from qualitative research? How should we exploit the new and emerging online methodology options? What extra dimensions can they bring to understanding and modelling choices?

## **2. DESIGN AND COST EFFICIENCY DEVELOPMENTS IN QUANTITATIVE DATA COLLECTION – KEEPING THE RELATIONSHIP DYNAMIC**

There is a growing requirement to improve the cost effectiveness of research components and this sits uneasily alongside a need for increased complexity.

A key development in enabling this has been the increased use of online methods for both qualitative (as discussed above) and quantitative methodologies. Indeed online approaches have come to be seen as the universal panacea by many.

Quantitative cost effectiveness can be improved via online opportunities. Indeed there has been a mass exodus from face-to-face approaches in recent years. And yet until quite recently most UK government departments stipulated that all research needed to be undertaken using a face-to-face methodology. But that was before austerity measures kicked in.

There are two main routes to online panels – purpose built and commercial. And there are, unfortunately, issues with both.

### **2.1 Purpose built panels**

In the case of purpose built panels, using, say, client-held customer email address files, the key issue is typically one of non-response levels.

Traditionally a paper self-completion approach with a response rate of 5–10 per cent would be rejected in favour of a telephone or face-to-face approach where a significantly higher level of response can be achieved. However it is seemingly more acceptable to the industry to accept similar 5–10 per cent response rates from an online self-completion approach. Why? Is it just that we are now more concerned about cost or is it that the technology of the approach has blinded us to the statistical realities?

## **2.2 Commercial online panels**

An alternative and very commonly used approach, is to buy into commercial online panels where the achieved response rate tends to be higher and can reach levels of 70–90 per cent. This approach is favoured by academics and now the public sector for its relative ease and cost effectiveness. However there are substantial limitations across a number of areas, in terms of penetration, in regard to cognitive issues and in relation to the ‘professionalism’ of respondents.

Each is discussed in turn.

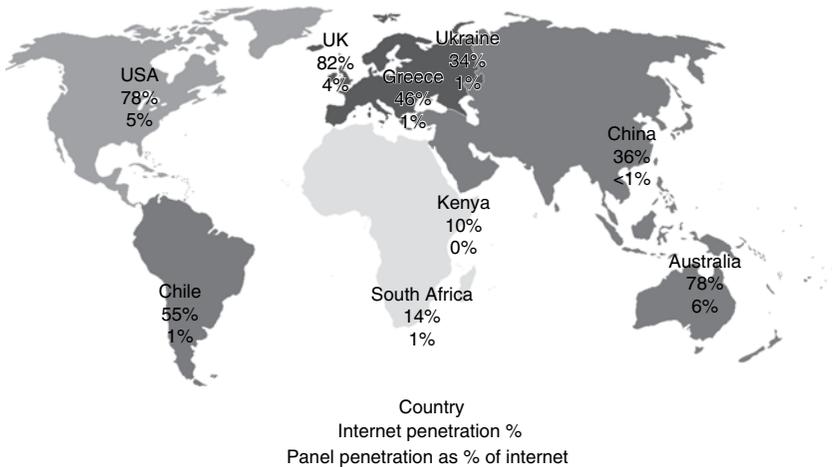
### **Penetration**

- there can be a scarcity of vulnerable respondents and by their very nature, many such respondents are unlikely to be recruited to a panel though some groups are better represented than others
- there is currently a lack of businesses on panels; respondents in the more senior roles within an organisation are unlikely to be able to pledge the time to an ongoing commitment
- there are also identifiable geographical weak spots, driven primarily by access limitations to broadband, thus skewing the sample away from these areas.

Most worrying is the penetration of the panel across the population base as a whole.

An illustration of the current internet panel penetration across a number of key countries is shown in Figure 6.1. This figure shows quite clearly that only a very small proportion of the population are being asked to respond to ongoing online research. In fact the penetration levels are probably worse than shown as many respondents belong to more than one panel and are thereby double counted.

This becomes a more worrying set of statistics when one takes into account the fact that many government organisations in the UK, for instance, are regularly polling the population on key policy decisions



Source: Data derived from 6 major panel providers.

Figure 6.1 Internet and panel penetration derived from the six major global panel providers

through this mechanism. And since most online panellists will respond to up to four surveys a week, these respondents are having a very large say on UK government policy.

### Cognitive issues

- when undertaking complex questioning, there is a lack of ongoing interviewer support to facilitate questionnaire completion and so in-depth probing is limited
- there is also no ability to obtain an interviewer rating about how well s/he thought that a particular respondent understood the tasks

### 'Professional' respondents

- as these respondents complete many surveys per week/per month, a crucial question arises as to what we know about their levels of survey/issue 'education' on the one hand and the impacts of survey fatigue on the other
- there are also many cited cases of respondents 'flat-lining' to respond quickly, i.e. without really looking at the questions properly and for example, just ticking the first box in the response options

### **2.3 Telephone**

So, face-to-face has been sidelined for reasons of cost and online has typically come in to replace it. This prompts the question as to why the industry seemingly chooses to largely ignore the telephone as a potential survey method. Many practitioners argue that it can be a very cost effective and robust option for both qualitative and quantitative research programmes. The many advantages of a telephone data collection approach are cited below. A major positive is that it can be truly random in a way that neither online nor face-to-face can be.

On the other hand there are emerging potential drawbacks to the method. Firstly the increased prevalence of mobile phones as the primary telephone contact number for many facilitates respondents' ability to 'screen out' any calls that they do not recognise, which can reduce response rates. Secondly there is the issue that mobile numbers are non location specific, which hinders any kind of geographical sampling process. Nevertheless there are clear benefits to using the telephone:

- it allows the full array of questionnaire customisation to be facilitated via automated question routing and filtering
- it can easily handle complexity: even when visual materials are required, these can be sent through to respondents by post or increasingly now online
- it is very time and cost effective
- it supports detailed diagnostic questions from both the interviewer and the respondent
- it enables interviewer supervision and support throughout the process.

So the question needs to be asked as to whether this approach has been mistakenly overlooked. Indeed, though there has been a rapid increase in the use of the telephone method within the market place, there still does not appear to be anywhere near full awareness or utilisation of the approach. This suggests that the approach requires further development and communication.

### **2.4 Customisation**

Another topic area covered concerned the issue of customisation in, for instance, stated preference surveys. Since the advent of stated preference approaches in the 1970s there has been a perceived value in customising the interview/stated preference context for respondents. The more a par-

ticipant feels that a particular survey is customised to their needs, activities, etc., the more able they seem to be to relate to the subject matter, and accordingly the more confident they are in responding/providing high quality data.

There are many examples in the stated preference discipline of how customisation can assist the respondent tasks and this has been facilitated by phone/internet approaches that allow for detailed customisation, for example of materials such as pictures, videos and comparison website mock ups, shown to the respondent whilst s/he talks to the interviewer on the phone.

There have also been some limited experiments undertaken to customise respondent location issues and material through GIS approaches, but concerns have been raised at different times about contravention of data protection issues if this is not carefully handled or thought through.

## **2.5 Cognitive overload**

The next issue addressed is a massive one for the industry and concerns the potential for cognitive overload. Understandably modellers want to eke as much information as they can from each and every respondent and there has been much practical support recently for the thesis that we market researchers have been too conservative in the past. The guiding principle that is emerging strongly is that if the subject matter is (made) interesting enough then respondents will stay focused and their cognitive abilities will be raised. But what is also important is the practice of including cognitive tests in advance of pilot testing of a survey tool to ensure that it is geared to respondents of all cognitive abilities.

## **3. IS THERE MORE TO COME THROUGH THE DEMANDS OF INCREASED MODELLING SOPHISTICATION?**

The key issue in the interpretation of all choice model data is the perfect confound between means and variances on the latent utility scale (Ben-Akiva and Lerman 1985). There are, literally, an infinite number of solutions and the analyst must use data that are external to the task, common sense and theory if he/she is going to identify the true solution.

A prime source of external data is the development of the choice experiment. Meta-ethnography and other new qualitative techniques have been used in development of attributes and levels for choice models (Al-Janabi et al. 2008, 2011). However the process of synthesising

qualitative data from individual studies to provide meta concepts also draws useful attention to the situations in which these concepts might manifest themselves differently in choice behaviour. It is also important to conduct such qualitative research within a multi-disciplinary team. This will help ensure that issues that are valuable to interpretation of the discrete choice experiments (DCE) (including the existence of dominant attributes) are recognised in attribute and level development (Coast et al., 2012).

Additional data from the DCE itself is perhaps where the most cutting-edge research is taking place. The inclusion of best-worst tasks, the relating of attitudinal statements to preferences and the splitting of choice tasks into those that elicit conditional and unconditional demand are three such examples (Flynn et al. 2007; Flynn 2010). Best-worst scaling (BWS) was invented by Jordan Louviere as a method of obtaining extra data in a cost-efficient way (Finn and Louviere 1992; Marley and Louviere 2005). It capitalises on the respondent's familiarity with an existing choice set, asking for worst or least preferred in addition to best or most preferred.

BWS can be used with three types of outcome and the 'Object Case' was recently used to quantify attitudes towards fifteen principles of health care reform in Australia (Louviere and Flynn 2010). These might traditionally have been elicited using category rating scales but BWS forced respondents to discriminate between them. Empirical distributions of attitudes can be plotted, showing the ability of BWS to test assumptions made by conventional models.

The second 'Profile Case' of BWS was used to value the ICECAP-O quality of life instrument. As for the Object Case, the best-minus-worst scores illustrate advantages of BWS, this time both in distinguishing a respondent with high choice consistency from one with low consistency and in identifying relative preferences for key attributes of quality of life (Flynn 2010).

In Table 6.1 the best-minus-worst scores are presented for each of four (ordered) levels for the five attributes of the ICECAP-O instrument. Each

*Table 6.1 Best-minus-worst scores for each of four (ordered) levels for the five attributes of the ICECAP-O instrument*

| ID | Attachment |    |   |   | Security |   |    |    | Role |   |    |   | Enjoyment |   |   |   | Control |    |   |   |
|----|------------|----|---|---|----------|---|----|----|------|---|----|---|-----------|---|---|---|---------|----|---|---|
|    | 1          | 2  | 3 | 4 | 1        | 2 | 3  | 4  | 1    | 2 | 3  | 4 | 1         | 2 | 3 | 4 | 1       | 2  | 3 | 4 |
| 1  | -2         | 0  | 2 | 2 | -3       | 0 | -1 | 3  | -1   | 0 | -1 | 1 | -1        | 0 | 1 | 2 | -3      | -2 | 0 | 3 |
| 2  | -4         | -1 | 4 | 4 | -1       | 0 | -1 | -1 | -2   | 1 | 1  | 2 | -2        | 0 | 1 | 0 | -3      | -1 | 1 | 2 |

attribute level was available to be picked four times across the (main effects) design. Thus the scores were bounded by  $\pm 4$ . Respondent ID#1 does not have a single 4 for any attribute level, whereas respondent ID#2 does (for attachment, which asks about love and friendship). Thus respondent #2 likely has a larger scale parameter (smaller variance) in his/her data. Individual values across levels for each attribute also provide insights into relative preferences: ID#1 values security and control whereas ID#2 values attachment. More complex models of preference and scale heterogeneity can be used, including those implemented by Latent Gold software (Flynn et al. 2010). However there is a continuing need to use theory and common sense to decide on the realism of the assumptions made by such models.

A second type of cutting edge data – those representing attitudes – have conventionally been quantified using rating scales. However Object Case BWS is now used to do so (Lee et al. 2007, 2008), capitalising on its invulnerability to person-specific response styles such as ‘yea-saying’ and avoiding extremes (Steenkamp and Baumgartner 1998). Indeed continuing use of rating scales risks throwing away the benefits of preference data by contaminating the choice based nature of the latter with measures whose mathematical properties are often suspect. Object Case BWS was used to avoid rating scale issues for one of the studies in the *Sydney Morning Herald* Independent Transport Enquiry. Three latent classes were found, all valuing improvements to “basic” services such as trains and buses, but two classes also valued integrated fares/ticketing and using green power for trains.

The final type of additional data concerns the use of structured choice tasks, particularly with reference to estimating both conditional and unconditional demand. A recent CenSoC, UTS housing study did this. Respondents first answered a conditional demand set of questions, indicating which dwelling they most preferred and which they least preferred. They were then asked if they would actually purchase their most preferred dwelling or remain in their current one. This allowed multiple models to be estimated, both conditional and unconditional, providing multiple insights in their utility functions.

Completion times are now routinely available for online surveys, both for the entire survey and increasingly for choice tasks within each choice set. These offer the potential to better understand variance heterogeneity as it manifests itself within the survey: it is already known that choice consistency typically improves during a ‘learning period’ in the first few choice sets, then remains constant until boredom sets in when it decreases again. When completion time data are equivocal, simple functions of best-worst data, particularly those elicited using Profile Case BWS, offer

insights into ‘click-through’ behaviour. These functions can be computed in real time, offering quality checks on web-survey data. This will have additional benefits to online panel providers in demonstrating data quality.

Eye-tracking is increasingly used to ascertain respondent attention to individual attributes. This offers both basic tests of whether they are ignoring attributes, and more complex insights into likely sources of respondent uncertainty; better modelling of within-person variance heterogeneity is likely to result. EEG data has been available for a number of years but it is only now that researchers are beginning to understand which components of the enormous sets of data it provides are useful in understanding preferences. It is likely that it will be used in conjunction with eye-tracking data, and CenSoC, UTS has already trialled this successfully.

Finally longitudinal (panel) data offer the prospect of understanding to what degree preferences and/or scales change over time. These data will provide a test of the predictions inherent in cross-sectional DCE data. Again being able to relate choice consistency to external variables (time) offers the ability to decompose the mean-variance confound that plagues DCE data.

#### 4. ADDITIONAL DISCUSSION ISSUES

Paul gave a short presentation which (re-)raised a number of issues that could form the basis for discussion. He split these into qualitative and quantitative areas. In summary he suggested that the main requirements and benefits from qualitative research comprised:

- looking at getting saturation and depth rather than representativeness
- to inform design issues
- to pre-test questionnaire and related instruments (i.e. a recommendation to make cognitives standard)
- to challenge and add depth to the quantitative findings.

An interesting question that he raised concerned whether we should be aiming to link qualitative research to choice experiment design and analysis, or whether we should be treating them as separate, complementary exercises. Paul suggested that, in many cases, the latter approach might be more appropriate given the difference in format between the focus group and the survey context.

With respect to quantitative research, Paul reinforced the need to

take a good strategic look at the data capture options available through face-to-face, online and phone. He reinforced the enhanced opportunities that exist now for customisation but asked whether there was a danger of this leading to endogenisation – when the design is no longer independent of the respondent's choices, then unobserved characteristics of the respondent could drive both the design and the choice, and thereby bias parameter estimates of the effect of the design.

Paul invited the workshop to consider whether there were tests or 'rules of thumb' that could be set to more systematically and simply explore issues of cognitive overload. Lastly he set out a wish list for the development and conduct of quantitative surveys which comprised:

- the use of computers as standard
- built in randomisation (ensuring orthogonality)
- each survey automatically tailored by location and by survey responses, linked to databases
- automatic dynamically optimised experimental designs (unique to each respondent)
- greater use of visuals, maybe even audio/video clips
- greater use of best-worst type questioning.

## 5. THE DISCUSSION

A robust discussion followed and a number of attendees talked about their experiences. Key themes that emerged from this were fed back to the conference. These comprised:

- a strong sense that there is a lot of untapped potential in data collection opportunities; there was a great deal of interest in the different options that had been presented during the earlier part of the workshop
- a need for data collection options to continue to keep abreast of the innovative thinking around modelling; but importantly also the view was strongly advocated that the understanding should also be continually sought in the opposite direction
- a strong endorsement that there should be continued development of mixed data collection methodologies as appropriate and that these needed to be monitored closely for performance
- an immense value was clearly attached to the development of cognitive tests but the question was repeatedly posed as to what was an acceptable level of potential sample cognitive acceptance (do we

design for 80 per cent of the population, or do we have to do so for 100 per cent at all times)?

- a strong scepticism from some about the use of professional panels for respondent recruitment, but, at the same time, a recognition of their strategic value
- an agreement that more care is needed in question framing which warrants an increased use of linguists and psychologists in research teams as a matter of course
- a need to explore in greater detail the need for customised language to account not just for national and cultural differences but also age segmentation, though it was recognised that these would equally put a strain on analytical compatibility
- there were some challenging initiatives raised in terms of deliberative approaches (e.g. to explore the potential for qualitative deliberative approaches to cross over into quantitative programmes) and in terms of 'social network' based research (e.g. getting a stated preference response from a social network peer group as a whole).

## REFERENCES

- Al-Janabi, H., J. Coast, and T.N. Flynn (2008), 'What do people value when they provide unpaid care for an older person? A meta-ethnography with interview follow-up', *Social Science & Medicine* **67**, 111–121.
- Al-Janabi, H., T.N. Flynn, and J.Coast (2011), 'Estimation of a preference-based Carer Experience Scale', *Medical Decision Making* **31**(3), 458–468.
- Anable, J. (2005), "'Complacent Car Addicts" or "Aspiring Environmentalists": identifying travel behaviour segments using attitude theory', *Transport Policy* **12**(1), January.
- Ben-Akiva, M. and S.R. Lerman (1985), *Discrete Choice Analysis: Theory and Application to Travel Demand*, Cambridge, MA: MIT Press.
- Bonsall, P., V. Stone, J. Stewart and M. Dix (2006), *Consumer Behaviour and Pricing Structures, Final Report on Qualitative Research to UK Department of Transport*, R104, January.
- Bonsall, P., Shires, J., Matthews, B., Maule, J. and Beale, J. (2007), 'Responses to complex pricing signals: Theory, evidences and implications for road pricing', *Transportation Research, Part A*, **41**(7), 672–683.
- Coast, J., H. Al-Janabi et al. (2012), 'Using qualitative methods for attribute development for discrete choice experiments: issues and recommendations', *Health Economics* **21**(6), 730–741.
- Crockett, J., G. Whelan, C. Sinclair and L. Hunt (2008), 'Willingness-to-pay for bus rapid transit: an examination of the influence of psychographics on choice', *European Transport Conference*, Leeuwenhorst Conference Centre, Netherlands.
- Finn, A. and J.J. Louviere (1992), 'Determining the appropriate response to

- evidence of public concern: the case of food safety', *Journal of Public Policy & Marketing* **11**(1), 12–25.
- Flynn, T.N. (2010), 'Valuing citizen and patient preferences in health: recent developments in three types of best-worst scaling', *Expert Review of Pharmacoeconomics & Outcomes Research* **10**(3), 259–267.
- Flynn, T.N., J.J. Louviere et al. (2007), 'Best-worst scaling: What it can do for health care research and how to do it', *Journal of Health Economics* **26**(1), 171–189.
- Flynn, T.N., J.J. Louviere, T.J. Peters and J. Coast (2010), 'Using discrete choice experiments to investigate heterogeneity in preferences for quality of life. Variance scale heterogeneity matters', *Social Science and Medicine* **70**, 1957–65.
- Hensher, D. (2009), 'Attribute processing, heuristics and preference construction in choice analysis', in A. Daly and S. Hess, *Choice Modelling: The State-of-the-Art and the State-of-Practice, Proceedings of Inaugural International Choice Modelling Conference*, Leeds, Emerald Publishing 2010.
- Le Masurier, P., J. Bates, J. Taylor, I. Flindell, D. Humpheson, C. Pownall and A. Woolley (2007), 'Attitudes to noise from aviation sources in England', Final Report for Department for Transport, MVA Consultancy Norwich, UK: Her Majesty's Stationery Office.
- Lee, J.A., G. Soutar et al. (2007), 'Measuring values using best-worst scaling: the LOV example', *Psychology & Marketing* **24**(12), 1043–1058.
- Lee, J.A., G. Soutar et al. (2008), 'The best-worst scaling approach: an alternative to Schwartz's values survey', *Journal of Personality Assessment* **90**(4), 335–347.
- Louviere, J.J. and T.N. Flynn (2010), 'Using best-worst scaling choice experiments to measure public perceptions and preferences for healthcare reform in Australia', *The Patient: Patient-Centered Outcomes Research* **3**(4), 275–283.
- Marley, A.A.J. and J.J. Louviere (2005), 'Some probabilistic models of best, worst, and best-worst choices', *Journal of Mathematical Psychology* **49**, 464–480.
- MVA Consultancy (2009), 'Understanding the passenger: valuation of overcrowding on rail services', Report to UK Department for Transport by MVA Consultancy in association with BAH and DeltaRail, December.
- Steenkamp, J.-B.E.M. and H. Baumgartner (1998), 'Assessing measurement invariance in cross-national consumer research', *Journal of Consumer Research* **25**(1), 78–90.