

How to make the most of online data collection technology

Paul Deakin
OPP Ltd

BPS Division of Occupational Psychology
Conference – 13th Jan 2012



What will we be covering?



- Why online data collection is topical
- Our project and concerns
- Best practice guidelines
- Personal insights
- Closing remarks
- Q&A

Why the subject is topical



- American Association for Public Opinion Research (AAPOR) Report on Online Panels (published in 2010)

“a decade ago the use of online surveys was virtually nonexistent.”

- Now very widely used
 - commercial organisations and governmental bodies
 - estimated spend: \$2 billion in 2009

Why the subject is topical



- Brought about by recent improvements in the availability and quality of online survey technology
 - Faster
 - Easier
 - Cheaper
- But at what cost?
 - to the integrity of the data
 - to the research participants themselves

Our project



- New 16PF research
 - 3 languages (English, French, Dutch)
 - Nationally representative samples
- 2 types of data:
 - 16PF (personality questionnaire)
 - Criterion data (to allow exploration of links between 16PF scores and real-world outcomes)
- Survey length: 1hr
- Volunteer online panel used



Our main concerns



- Sampling
 - national representativeness, across several variables
- Quality of data collected
 - Conscientious responding?
 - Respondent identity?
 - Response format impact (ie online)?

Best practice guidance



Best practice



- Based on ESOMAR* guidelines
- 4 over-riding principles:
 1. treat respondents with respect
 2. be sensitive to consumer concerns (& public confidence)
 3. maintain a distinction between research and commercial activities
 4. adhere to data protection requirements

**ESOMAR – European Society for Market Research*

Areas covered



Company profile

Sample source

Panel recruitment

Panel and sample management

Policies and compliance

Partnerships and multiple panel partnership

Data quality and validation

Company profile



- What experience does the company have with providing online samples for market research?

- What is the source(s) for the online sample?
databases, actively managed panels, direct marketing lists, web intercept sampling, or other
 - Why is this source the most appropriate?
 - If source is a panel or database, is this used solely for market research?
 - How are hard-to-reach groups sourced?
 - What are people told when they are recruited?

If the sample comes from a panel:

- What is the annual panel turnover /attrition / retention rate?
- What is the opt-in process?
- How is respondent identity ensured?
- What profile data is kept on panel members?
- What is the size and/or the capacity of the panel?

Panel and sample management



- What is the sampling process?
- How are people invited to take part in a survey?
- How are respondents incentivised?
- How often are individual members contacted for online surveys within a given time period?

- What privacy policy is applied?
 - Compliant with all regional, national and local laws with respect to privacy, data protection and children?
- What data protection/security measures are in place?
- What quality management system is applied?
- If surveying children, how is permission obtained?

Partnerships and multiple panel membership



- Are samples supplemented with samples from other providers?
- Is there a policy regarding multi-panel membership?

- What are likely survey start rates, drop-out and participation rates? How are these computed?
- Is individual level data such as recent participation history, date of entry, source, etc., stored on panellists?
- What data quality analysis and validation techniques are used to identify inattentive and fraudulent respondents?
- Is respondent satisfaction measured?
- What information is provided to debrief clients after the project has finished?

Personal insights



What have we learnt?



The positives

- Lots of competing providers out there (with widely differing prices)
- Geared towards collecting data quickly
- It is reasonably cost-effective (e.g. compared with opportunistic data collection)

What have we learnt?



Challenges

- Short surveys the norm (<20 mins)
- Providers cagey about disclosing exactly how they ensure 'accurate' survey responding
- Difficulty obtaining truly representative sampling

Closing remarks



- Probability-methods may be better – eg Census approach
 - but, volunteer online panels do have their uses (fit for purpose)
 - No conclusive research on how good different approaches are
- Different panels/providers have different standards
 - eg Caution when claiming national representativeness
- Panel providers could help the debate by sharing more about their methods
- Test publishers should
 - disclose the methods used
 - Conduct additional quality checks on the data themselves

Thank you

Any questions?

paul.deakin@opp.eu.com
