

ESRC Longitudinal Studies Review 2017

Further analysis of responses to the consultation

Paper 10:

Representativeness and study design

April 2017

Report author:

Ruth Townsley, Independent Researcher

The views represented in this report are from those who responded to the consultation and do not represent the views of ESRC

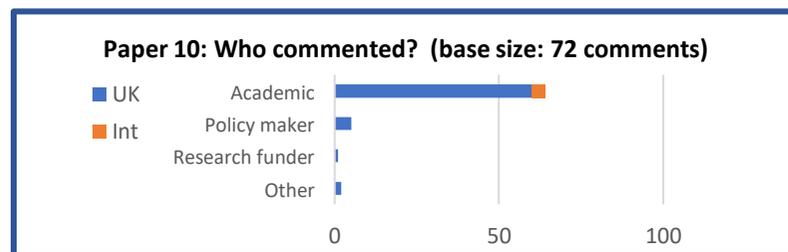


Introduction

The ESRC Longitudinal Studies Review 2017 is exploring the current and future scientific and policy-relevant need for longitudinal research resources. The review started in October 2016 and will report to ESRC Council early in 2018. An open online consultation in autumn 2016 sought input broadly, resulting in 637 completed responses from UK (83.4%) and international (16.6%) respondents. Respondents were predominantly from the academic sector (80%) as well as government, civil society and business sectors (20%). The main findings of the consultation were published in December 2016 in an [initial report](#). This report is supplemented by short briefing papers that examine key themes from the consultation data in more detail.

Paper 10: Representativeness and study design

This paper examines the comments made by consultation respondents about representativeness as it relates to study design. There were 72 comments made about this topic, identified via answers to question 9 that were coded as 'sampling and population representation' and via a whole-survey search using the strings representat* and probability.



"... it is vital for social scientists, and for this review, to re-state the importance of population representation based on probability sampling for the UK's longitudinal studies. The MRC's 2014 review does not discuss probability sampling at all, and although biomedical researchers are also concerned about representation and generalisability their inferential framework is a very different one." (ID 688)

The importance of population representation

Respondents were unanimous in advocating for the continued centrality, and funding, of national population representation in longitudinal studies and its foundation in probability-based sampling. Respondents highlighted that the nationally representative nature of the studies' samples sets them apart from 'big data' and from other non-representative data collections, and that this offers significant scientific advantages:

- > National representative samples allow for robust evidence to be produced, are the only way to accurately compare cohort data and are vital for understanding socio-economic inequalities and mobilities
- > They enable effective evaluation of the impact of policies on different domains of people's lives and different subgroups of the population, an area of research which looks set to increase since devolution and Brexit
- > Population representation is at the core of economic and social science, and the survey method relies on statistical inference from samples to populations for generalisability - this cannot be achieved by using administrative datasets or other forms of big data, whose findings will continue to need validation against large-scale representative population surveys.

A few respondents argued strongly that a continued commitment to supporting probability-based population representation as central to longitudinal studies should be made explicit by ESRC through the 2017 review, particularly within the context of the merger of the UK research councils into UKRI.

"The cohort studies are a prime resource for understanding socio-economic inequalities... Representative social survey data is vital in representing these groups who are typically very poorly represented in non-representative samples." (ID 126)

Maintaining representativeness

The need to minimise attrition and non-response in order maintain representativeness was a key theme, mentioned by many respondents. The issues covered included:

- > Keep selective attrition to a minimum, by using all available technology and methodological approaches
- > Ensure continuing representation through refresher samples and adding new immigrants
- > Consider including sample weights to correct for missing data
- > Understanding of how mixed mode design may have implications for representativeness
- > Consider the use of administration and big data to help fill gaps.

"The importance of probability sampling in current and future cohorts ... is highlighted by the increasing use of "Big Data"... which implicitly assume that sample characteristics and estimated associations are transferable to the general population. Considering that this assumption is unverifiable it's of vital importance that longitudinal surveys remain representative of their target population as they can provide reliable insights on topics not covered by administrative and other forms of "Big Data", but also act like a gold standard where the generalisability of findings from "Big Data" and other unbalanced samples can be empirically verified." (ID 81)

"There are challenges to representativeness of both longitudinal studies (attrition and non-response) and big data including both government and commercial (significant gaps in coverage and hence difficulty in generalising to the population). However, combining the two provides significant opportunities to assess the impacts on population inference of these challenges, and hence provides an important opportunity to strengthen both." (ID 54)

Is there a need for a web-based probability panel?

Five people mentioned that the review should consider the case for funding a national web-based probability panel for the following reasons:

- > Cost-effectiveness when set against face-to-face fieldwork costs
- > Potential for more frequent waves and speed of turnaround
- > More regular, easier contact with participants thus less attrition and non-response
- > It represents a gap in the UK's longitudinal portfolio - other countries are successfully using web-based probability panels to address questions relating to short-term longitudinal change.

"Web-based probability panel - Face-to-face surveys are enormously expensive and are usually only possible with large time-intervals between waves. Online-based surveys can be done on a more regular basis and it is easier to follow people as they are more and more mobile. Similar studies exist in the Netherlands (www.lissdata.nl/lissdata/), Germany (http://reforms.uni-mannheim.de/internet_panel/home/) and France (www.elipss.fr/)... I would welcome the inclusion of innovation samples to allow for experimental work to be carried out as part of the panel." (ID 245)

"One possible issue that the review may wish to cover is whether or not the UK's longitudinal portfolio would benefit from a probability-based web panel. I recently co-authored a report on this topic for the ESRC. There is a perceived gap in the portfolio in relation to short-term longitudinal change which may this may be able to address. It will be interesting if this is a theme in the review and consultation." (ID 688)

Other panel issues

A range of other panel-related issues were mentioned which included:

- > Including siblings and twins in future studies
- > Considering more family-based designs, which include following up post-separation parents and the adult children of respondents
- > Including data on deceased individuals
- > Maintaining contact with emigrants.

"Many hypotheses can only be meaningfully answered using family based designs i.e. trios of mother-father-offspring data, or sets of siblings. Future rounds of data collection should collect more data on family members... Interviewing the (adult) children of respondents, would provide immense information about intergenerational/genetic links and transfers." (ID 596)

"Siblings and twins please! Including siblings and twins in future studies would help a great deal in terms of inferring causality from studies using the UK cohort data. The data is already so fantastic in many respects that it's a shame we must end every research paper with the limitations that our findings are 'only correlational'." (ID 337)

"Emigrants should not be 'attrition'! Instead of giving up on people who are known to have emigrated, I think they should be retained in the survey, at least if it's possible to do the interview (e.g. by telephone, or self-completion by post or internet). I know they are no longer seen as part of the "British population", but I think that's an overly rigid view. A "transnationalism" perspective on migration makes it clear that emigrants can retain attachment and influence in the country they leave. There are good reasons to make an effort to retain them as part of the sample. This is something the GSEOP is trying to do as well." (ID 371)