

thinking

Labels, damn labels and statistics

Sara Davidson, Ipsos MORI Scotland

The Scottish Government has been working to establish a collaborative strategic framework to facilitate increased cross-sectoral data linkages for research and statistical purposes. These linkages involve joining datasets to significantly increase the power of analysis possible with the data. One of the key findings to emerge from our deliberative research among members of the public was that this leads to the crude categorisation or “labelling” of individuals and groups.

In the spring 2009 edition of our newsletter [Approach](#), my colleague, Steven Hope, suggested that there is scope for public bodies to make more use of existing statistical and research evidence. Similar sentiments have been expressed, with increasing regularity, by public bodies themselves over recent years. Indeed, the Scottish Government has been working to establish a collaborative strategic framework to facilitate increased cross-sectoral data linkages for research and statistical purposes. These linkages involve joining two or more administrative or survey datasets and have the potential to significantly increase the power of analysis possible with the data, reducing the need for additional data collection. The Strategy incorporates the Research On Census Alternatives (Beyond 2011) project, an ongoing investigation of administrative data (such as data from electoral registration, state schools, DWP customer lists etc.) which may help produce population statistics without the high cost of a census.

Earlier this year the Government launched a consultation on the Data Linkage Framework. It also commissioned Ipsos MORI, along with staff from the Centre for Population Health Sciences at the University of Edinburgh, to undertake a series of public deliberative workshops to provide better evidence on the public acceptability of data linkage. The findings from both exercises have now been

published and can be accessed on the Scottish Government website¹.

One of the most interesting and unexpected findings to emerge from the deliberative workshops concerned participants' attitudes, not to data linkage, but to the general focus on quantitative data in decision making. There was a view that this leads to the crude categorisation or “labelling” of individuals and groups – for example, as being ‘from a bad area’ or ‘low achieving’ or ‘criminal’ – and subsequently to stigmatisation and discrimination. A reverse effect was also identified whereby individuals or groups who have *not* been labelled or categorised in a particular way miss out on much needed support or assistance as a result – the example was given of a small impoverished area not receiving financial assistance from government simply because it is not officially classified as one of the most deprived places in the country.

There was some concern that data linkage could exacerbate these problems by creating the potential for labels to carry across sector boundaries and receive wider application. A specific concern was that someone's past involvement with the criminal justice system could become known to various authorities

¹ [‘A Scotland-wide Data Linkage Framework for Statistics and Research: Consultation Analysis’](#).
[‘The Public Acceptability of Cross-sectoral data linkages: deliberative research findings’](#).

and result in them being placed at the bottom of a housing list or otherwise facing unequal access to services.

These concerns about linkage are largely unfounded as the Strategy is primarily concerned with linking *anonymised* data for research and statistical purposes, *not* sharing *personal information about an individual* between organisations. When the workshop participants were reassured on this point, most immediately became more comfortable with the idea of linkage. Still, their broader concerns about the potentially negative impact of categorising individuals and groups cannot be so easily negated and provide two important reminders to those of us working in

social research and policy. The first concerns the inherent limitations of aggregate – and indeed much sub-aggregate –level data analysis in promoting an understanding of individuals' lives, and the importance of remaining alert to atypical patterns of experience and need. The second is the considerable capacity of the public to engage at a sophisticated level with complex policy debate, and to shape that debate by drawing attention to, and questioning, taken-for-granted assumptions and practices on the part of decision makers.

Acknowledgements and copyright

This thinking is based on analysis undertaken by Ipsos MORI on its own initiative based on a research project conducted by Ipsos MORI on behalf of the Scottish Government. The full report can be accessed via the following link <http://www.scotland.gov.uk/Topics/Statistics/datalinkageframework>

For more information about this work contact

Sara Davidson
Associate Director
Ipsos MORI Scotland
4 Wemyss Place, Edinburgh, EH3 6DH

sara.davidson@ipsos.com
0131 226 8673

www.ipsos-mori.com/scotland