

# Why your research may be out of order

**A**RE older people more concerned about crime than young people? Do our political attitudes become more conservative with age? Do older generations feel differently about abortion than younger generations? How do our consumption patterns change across the lifespan?

Much of what we know about such issues is based on self-reports gathered in surveys. Unfortunately, the answers we obtain are, in part, shaped by the specific survey questions we ask. As researchers are well aware, minor changes in question wording, question order or response format can result in major differences in the reported behaviours or opinions (e.g. Schuman & Presser, 1981). Thanks to two decades of research at the interface of psychology and survey methodology, the cognitive and communicative processes underlying self-reports are by now reasonably well understood (for reviews see Schwarz, 1999b; Sudman *et al.*, 1996; Tourangeau *et al.*, 2000). What has received little attention, however, is how these processes are affected by age-related differences in cognitive functioning.

As Western societies age, the information needs about the life conditions, consumer behaviours, attitudes and routines of older people are steadily increasing. Many surveys are now conducted with elderly respondents or include a substantial proportion of them. Recent research has shown that younger and older people are differentially affected by features of the research instrument (see Schwarz *et al.*, 1999). Such differential influences can lead researchers to draw dramatically different conclusions about age differences in opinions and behaviours,



**BÄRBEL KNÄUPER and NORBERT SCHWARZ** explain how age-sensitive context effects may lead us astray.

depending on the way in which the specific question is asked. This article highlights some of the relevant findings.

## Age and response-order effects

Normal human ageing is accompanied by profound cognitive changes, ranging from decreased sensory functioning to the slower execution of cognitive processes and a general decline in working memory capacity (for an overview see Park, 2000). Not surprisingly, these changes influence the cognitive processes underlying self-reports.

Suppose, for example, that an interviewer asks you the following question over the phone: 'What is the most important problem facing our country today? The environment, taxes, unemployment, European unification, or national security?' Ideally, you listen to the full question, keep all problems in mind, think about what makes each one of them important, and then compare them to determine which you consider most important. This is a demanding task, in particular under the time pressure of telephone interviews, where question-answer sequences rarely exceed one minute. Under these conditions, respondents are more likely to elaborate on the last few alternatives read to them, which are still in their ear. When these alternatives are agreeable, they may be chosen without much consideration of the ones read earlier. This gives rise to recency effects in telephone interviews: an alternative is more likely to be chosen when presented late rather than early in the list. On the other hand, when the alternatives are presented in a written questionnaire, respondents can think about them at their own pace and in the order in which they are presented (instead of having

to listen to an interviewer, who moves on to the next alternative). When an early response alternative is agreeable, it may be chosen without much thought to the later ones. This gives rise to primacy effects when the question is presented in writing: an alternative is more likely to be chosen when presented early rather than late in the list (for a discussion see Krosnick & Alwin, 1987; Sudman *et al.*, 1996).

Given older adults' slower processing speed and reduced memory capacity, we predicted that older respondents would be particularly likely to show response-order effects in telephone interviews. We looked at studies where depending on conditions, a response option was read out as either the first or last alternative. We expected that older respondents would be more likely to agree with a given alternative when it was read out last. A meta-analysis of these response-order experiments in large-scale surveys confirmed our prediction (Knäuper, 1999b), with an average recency effect across 14 experiments of 31 percentage points for respondents aged 65 and older, but only 14 percentage points for respondents younger than 65.

Figure 1 shows a particularly dramatic example. In this study (Schuman & Presser, 1981) respondents were asked in a telephone interview, 'Should divorce in this country be easier to obtain, more difficult to obtain, or stay as it is now?' The response alternative 'more difficult' was read to respondents as either the second or the last alternative. Consistent with the general results of the meta-analysis, the size of the observed recency effects increased with respondents' age, ranging from a nonsignificant difference of 5 percentage points for age 54 and younger to a whopping 36.4 percentage points for age 70 and older.

## WEBLINKS

Survey Methodology Program, Institute for Social

Research: [www.isr.umich.edu/src/smp](http://www.isr.umich.edu/src/smp)

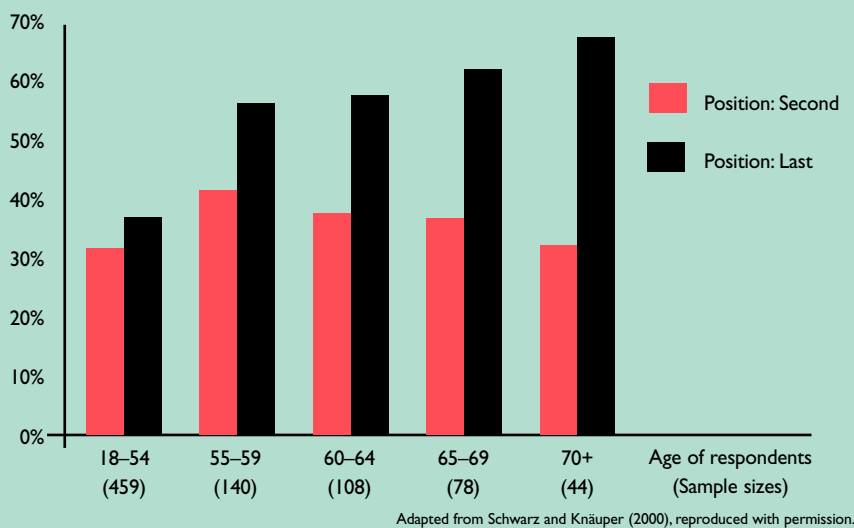
Norbert Schwarz's homepage:

[sitemaker.umich.edu/norbert.schwarz](http://sitemaker.umich.edu/norbert.schwarz)

Bärbel Knäuper's homepage:

[www.psych.mcgill.ca/perp/gfac/knaeuper](http://www.psych.mcgill.ca/perp/gfac/knaeuper)

**FIGURE 1** Percentage of respondents endorsing the survey option that divorce should be more difficult to obtain, according to whether the option was given as the second or last option in a list



Note also that we would draw different substantive conclusions about the relationship of age and attitudes towards divorce, depending on the order in which the response alternatives are presented. Whereas attitudes towards divorce seem to become much more conservative with age when 'more difficult' is the last response option, no reliable age differences emerge

under the other order condition. As this example illustrates, age-sensitive context effects can invite misleading conclusions about the phenomenon being studied.

#### Age and question-order effects

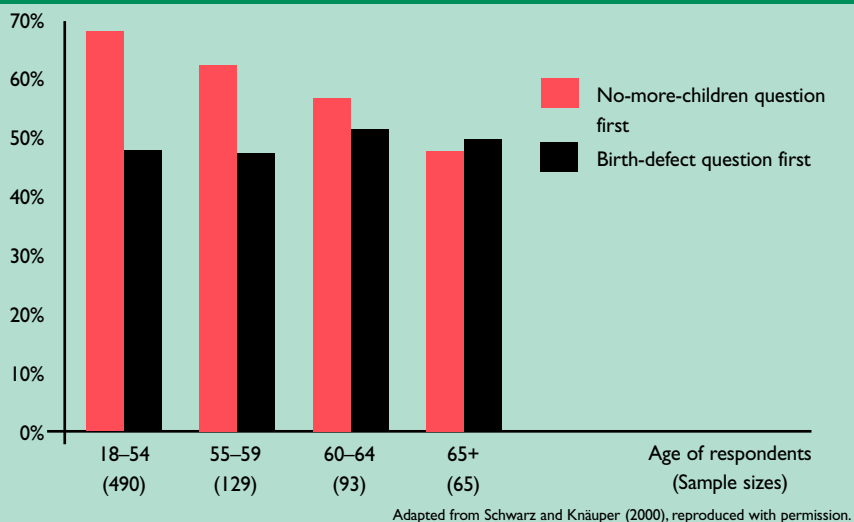
Numerous studies have demonstrated that preceding questions can profoundly influence respondents' answers to

subsequent questions (e.g. Schuman & Presser, 1981). Such question-order effects emerge when a question brings information to mind that respondents may otherwise not consider in answering the next one (Sudman *et al.*, 1996). Ironically, however, what results in larger response-order effects in older age should reduce the size of question-order effects. Much as age-related declines in working-memory capacity make it more difficult to keep all response alternatives in mind along with their pros and cons, they may also make it less likely that information used to answer an earlier question is still accessible in memory when a related question is asked later on. If so, question-order effects should be attenuated in older age. Again, the available data support this prediction (Knäuper, 1999a; Schwarz & Knäuper, 2000).

As an example, consider one of the most robust question-order effects in the survey literature. Schuman and Presser (1981) asked respondents two questions about abortion, namely whether a pregnant woman should be able to obtain a legal abortion 'if she is married and does not want any more children' (Question A) or 'if there is a strong chance of a serious defect in the baby' (Question B). In general, respondents are more likely to support legal abortion in response to the 'birth-defect question (B) than in response to the no-more-children question (A). More importantly, the endorsement of the no-more-children question drops when people are first asked the birth-defect question: compared with the risk of a serious birth defect, merely 'not wanting any more children' appears less legitimate, reducing support for legal abortion. Figure 2 shows how this question-order effect changes with age. While younger respondents show a question-order effect of 19.5 percentage points, the effect decreases with age and is no longer reliably obtained for the over-60 age group.

Once again, we would arrive at different conclusions about age differences in attitudes towards abortion, depending on the order in which the questions were asked. When the no-more-children question is asked first, support for abortion decreases with age, suggesting that older respondents hold more conservative attitudes towards abortion. Yet attitudes towards abortion are thoroughly independent of age when the same question is preceded by the birth-

**FIGURE 2** Percentage of respondents supporting abortion for married women who do not want any more children, according to whether the survey question preceded or followed a question about support for abortion to avoid birth defects

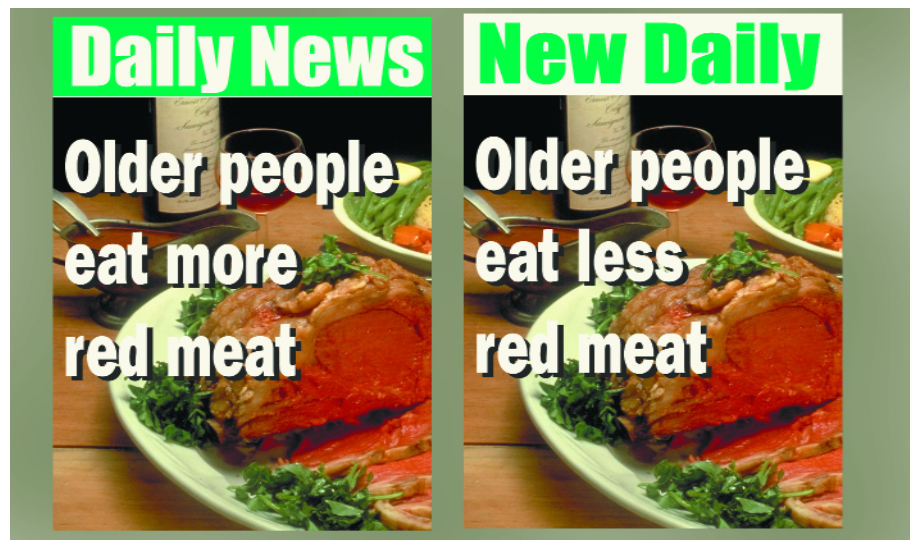


defect question. Hence, the order in which a researcher happens to present two related questions can determine whether age differences in attitudes will emerge.

### Behavioural frequency reports

Surveys often ask how frequently one does certain things, for example, how often one has bought coffee in the past four weeks. Such events are usually not well represented in autobiographical memory. Because people have a hard time recalling the frequency of mundane behaviours and events, they typically rely on estimation strategies (for a review see Schwarz & Oyserman, 2001). One such strategy draws on the frequency scale presented by the researcher. In a nutshell, respondents assume that the researcher constructs a meaningful scale. Hence, values in the middle range of the scale presumably reflect the 'average' or 'typical' behaviour, whereas values at the extremes of the scale correspond to the extremes of the distribution. Based on this assumption, respondents use the frequency scale as a frame of reference in estimating the frequency of their own behaviours. This results in higher frequency reports along scales with high rather than low values. For example, 37.5 per cent of a sample of German respondents reported watching TV for more than two hours a day when given a high frequency scale ranging from 'up to two hours' to 'more than four hours' a day. In contrast, only 16.2 per cent reported doing so when given a low-frequency scale ranging from 'up to an hour' to 'more than two hours' (Schwarz *et al.*, 1985). Such scale-based estimation effects are more pronounced the more poorly the behaviour is represented in memory (Menon *et al.*, 1995).

It should be even harder for older people to remember how often they did something. Thus, we may assume that they are particularly likely to resort to estimation strategies and may therefore be more influenced by the frequency scale presented to them. Indeed, this has been found (e.g. Schwarz, 1999a). In one experiment, 24 per cent of the young respondents (aged 29–40) reported 'eating red meat' 10 times a month or more when presented with a low-frequency scale, whereas 43 per cent did so when presented with a high-frequency scale – a difference of 19 percentage points. In contrast, 19 per



Whether a high- or low-frequency scale is used can have a major impact on findings

cent of older adults (aged 60–90) reported this frequency when presented with a low-frequency scale, but 63 per cent did so when presented with a high-frequency scale – a difference of 44 percentage points.

Again, note that if you research age effects this finding has major implications for you. We would arrive at different conclusions about age differences in the consumption of red meat, depending on the scale the researcher happened to use: the answers in response to the low-frequency scale suggest that young people eat somewhat more red meat than older people (a nonsignificant difference of 5 percentage points), whereas the answers in response to the high-frequency question suggest that older people eat considerably more red meat than younger people (a reliable difference of 20 percentage points).

While older people may not recall well how often they ate red meat, there are other behaviours and experiences to which they are likely to pay close attention. Specifically, older people are known to monitor their health and physical symptoms more closely than young people (e.g. Deeg *et al.*, 1996). If so, this information should be better represented in the memories of older than of younger respondents, reversing the previously observed influence. Supporting this conjecture, we found that older respondents were less susceptible to frequency-scale effects for questions asking about physical symptoms. For example, 37 per cent of younger adults reported having headaches 'twice a month or more often' when

presented with a low-frequency scale, whereas 56 per cent did so when presented with a high-frequency scale – a difference of 19 percentage points. In contrast, older adults' frequency reports were virtually unaffected, with 11 per cent reporting this frequency along a low-frequency scale, and 10 per cent along a high-frequency scale (Schwarz, 1999a).

These examples again illustrate that older people are in certain cases more, and in others less, vulnerable to response effects in surveys. The quality of survey data does not generally decrease with increasing age of the respondents – whether the responses of older people will be more or less biased than those of younger people depends on the formal features and content of the scale used.

### What to do about it?

When age differences in attitudes or behaviours are observed – unexpectedly or expectedly – one is well advised to carefully examine whether they could be a methodological artefact. As the reviewed examples indicate, any observed difference may reflect a true difference in the attitudes or behaviours being studied, a differential influence of features of the research instrument, or an unknown mix of both. 'Age-sensitive' questionnaire design requires that researchers think through the issues at hand for every question asked and minimise the differential impact of memory on survey responses. For example, they may want to show the response options on show cards, use filler items between

substantively related questions, or replace numeric frequency scales with open-ended frequency questions (for general guidelines see Sudman *et al.*, 1996; Schwarz & Oyserman, 2001).

Possibly, the next time you see a news headline saying 'Poll shows that older people are more opposed to human cloning', you might wish that the journalist had provided more details on how exactly the question was asked, including the response options used, the order in which they were presented, and the content of preceding questions. Maybe there is more agreement between the generations than the reported results indicate – if only the question were asked in the right way.

■ *Dr Bärbel Knäuper is at McGill University, Montreal. E-mail: barbel.knauper@mcgill.ca.*

■ *Professor Norbert Schwarz is at the University of Michigan. E-mail: nschwarz@umich.edu.*

## References

- Baddeley, A.D. (1986). *Working memory*. Oxford: Clarendon Press.
- Deeg, D.J.H., Kardaum, J.W.P. & Fozard, J.L. (1996). Health, behavior, and aging. In J.E. Birren & K.W. Schaie (Eds.) *Handbook of the psychology of aging* (pp.129–149). San Diego: Academic Press.
- Knäuper, B. (1999a). Age differences in question and response order effects. In N. Schwarz, D. Park, B. Knäuper & S. Sudman (Eds.) *Cognition, aging, and self-reports* (pp.341–363). Philadelphia: Psychology Press.
- Knäuper, B. (1999b). The impact of age and education on response order effects in attitude measurement. *Public Opinion Quarterly*, 63, 347–370.
- Krosnick, J.A. & Alwin, D.F. (1987). An evaluation of a cognitive theory of response order effects in survey measurement. *Public Opinion Quarterly*, 51, 201–219.
- Menon, G., Raghurir, P. & Schwarz, N. (1995). Behavioral frequency judgments: An accessibility–diagnosticity framework. *Journal of Consumer Research*, 22, 212–228.
- Park, D.C. (2000). The basic mechanisms accounting for age-related decline in cognitive function. In D.C. Park & N. Schwarz (Eds.) *Cognitive aging: A primer* (pp.3–22). Philadelphia: Psychology Press.
- Schuman, H. & Presser, S. (1981). *Questions and answers in attitude surveys*. New York: Academic Press.
- Schwarz, N. (1999a). Frequency reports of physical symptoms and health behaviors: How the questionnaire determines the results. In D.C. Park, R.W. Morrell & K. Shifren (Eds.) *Processing medical information in aging patients: Cognitive and human factors perspectives* (pp.93–108). Mahwah, NJ: Lawrence Erlbaum.
- Schwarz, N. (1999b). Self-reports: How the questions shape the answers. *American Psychologist*, 54, 93–105.
- Schwarz, N., Hippler, H.J., Deutsch, B. & Strack, F. (1985). Response categories: Effects on behavioral reports and comparative judgments. *Public Opinion Quarterly*, 49, 388–395.
- Schwarz, N. & Knäuper, B. (2000). Cognitive aging and accuracy of self-report. In D. Park & N. Schwarz (Eds.) *Cognitive aging: A primer* (pp.233–252). Philadelphia: Psychology Press.
- Schwarz, N. & Oyserman, D. (2001). Asking questions about behavior: Cognition, communication and questionnaire construction. *American Journal of Evaluation*, 22, 127–160.
- Schwarz, N., Park, D., Knäuper, B. & Sudman, S. (Eds.) (1999). *Cognition, aging, and self-reports*. Philadelphia: Psychology Press.
- Sudman, S., Bradburn, N.M. & Schwarz, N. (1996). *Thinking about answers: The application of cognitive processes to survey methodology*. San Francisco: Jossey-Bass.
- Tourangeau, R., Rips, L.J. & Rasinski, K. (2000). *The psychology of survey response*. Cambridge: Cambridge University Press.