2 Concepts and Topics

Mathis Schröder

2.1 Introduction

As mentioned in the introduction to this book, one of the principal difficulties faced by longitudinal surveys such as the Survey of Health, Ageing and Retirement in Europe (SHARE) concerns how to deal with "initial conditions" – the lives of respondents before the baseline year of a survey. Due to the lack of better options, the only realistic way for retrieving past information in certain domains is to obtain the information directly from the respondents. However, people do not reproduce events from the past flawlessly (Rubin, 1996), and the characteristics of the individual, the type of data collected and the period of recall all influence the accuracy of recall (Rubin and Baddeley, 1989; Sudman and Bradburn, 1973). An improved understanding of the nature of memory has been important in developing data collection techniques that improve accuracy of recall. Belli (1998) explains that "memories are thematically and temporally structured within a hierarchical ordering".

In the following sections, this chapter documents the path from the literature on recall error and ways to improve recall in data collection efforts to the point of the electronic implementation for the SHARELIFE study.

2.2 Recall error in retrospective data collection – some evidence

In any survey the main issue of answering questions, retrospective or not, is accuracy. Most socio-economic surveys have a retrospective component to them, for example when asking about events that happened in the last year or when collecting information on the frequencies of events. That there is room to improve accuracy in retrospective surveys has been shown in many studies across different subjects. In the following, several examples of known issues are illustrated and briefly discussed.

Peters (1988) uses the National Longitudinal Survey of Work Experience (NLS), a panel that started in 1966. As marital status – and thus changes in it – is recorded each year, she can construct a marriage history for her study group of young women. In 1978 and in 1983, the women are asked about their last three marriages, which gives the opportunity to compare the retrospectively collected data with the ongoing records of marital history. Taking the "at-the-time" recordings as the true data, Peters finds a concordance of between the panel and the retrospective data of 76%. The underlying recall period is at the maximum fifteen years, with a mean distance between interview and event of slightly more than five years. When looking at the determinants of this recall error, she finds that the recall is reduced by 3-4% per year. In addition, education is reducing recall error.

Auriat (2000) studies moving histories. The retrospective data source is a survey of couples in Belgium, who are asked separately and as a couple about their residential changes from when they were 14 years old. As they are between 41 and 57 years old at the time of the interview, there are a potentially large number of moves to report. The comparison data come from the Belgium National Population Register, where all moves have to be reported to within 8 days of changing the address. Auriat sets the error margin to three months: if the reported move date is within three months of the official move date, there is no error. Considering at most the first three moves after marriage, she finds that at least 30% are off by more than 3 months. Quite interestingly, however, she finds that if the reports are wrong, respondents miss the mark by yearly amounts, i.e. the difference to the true date is by multiples of 12 months rather than by any other difference. This is attributed to memories of the season, in which the move happened. She finds little evidence that time since event matters, which might be an effect of the moves all being long ago.

Mathiowitz and Duncan (1988) use the PSID in 1983 to assess the accuracy of retrospectively collected unemployment information. The respondents are asked about their monthly unemployment history in the previous two calendar years. From this, an indicator of "any unemployment" is generated in combination with the detailed unemployment history up to the interview. The employment records of a large firm are used to validate this information. A look at the indicator yields corresponding answers in 89% of the data – i.e., 89% of respondents indeed had an unemployment period if they reported it in the PSID and vice versa. However, the duration of unemployment is often misreported: specifically spells with a short duration of 12 weeks or less are omitted by the respondents in more than 50% of the time, and still 37% of spells that last 29 weeks or longer are not reported. Interestingly, time since the event is not related to the recall error, maybe an effect of the short time period considered here.

A study related to health by Means et al (1989) was done for the US department of Health and Human Services. They collect data on doctor visits for any health condition over the last 12 months, and compare the reported visits to those in the actual medical records. Only about 41% of all visits are recalled, and, interestingly, the seriousness of the condition does not matter in terms of recall. Recurring events, i.e. those where the respondent had to come back for the same condition, were recalled with less precision: visits for the same condition that led to three or more visits are each less accurately recalled than less frequent ones, with a margin of 21 percentage points, or 32 vs. 53%.

There has also been an effort to test the memory of public events, done by Gaskell et al (2000) in the United Kingdom. In this study, people were asked to remember two public events: Margaret Thatcher's resignation, which at the time of the interview had been 19 months ago, and the Hillsborough football disaster, 37 months ago at the time of the interview. Correct responses were relatively few, only 15% of respondents recalled the exact year and month of the resignation, whereas 10% were able to exactly date the football disaster. Given the previous examples, this hints at the possibility, that public events are less accurately remembered than events in respondents' personal histories.

These examples show that there is evidence of recall bias in various topics, although there are differences among the topics in terms of how easy things are forgotten. Personal moves are much better recalled than visits to the doctor, for example, especially when considering the differences in time horizon in these studies. This leads to the very intuitive point that events that are more important in a respondent's life are more easily remembered.

This hypothesis of a saliency effect is tested in two studies by Akerlof and Yellen (1985) and Jürges (2007), which both look at the correct report of unemployment spells that had happened a year prior to the interview, dependent on the impact this had at the time on the respondent's life. While Akerlof and Yellen measure the impact indirectly, Jürges uses the reported life satisfaction to explain recall bias. Both studies find that the more important the event was at the time, the more accurate it is remembered later. However, this is of little practical purpose, as it is unknown to the researcher, what are the salient events in a person's life.

Given this brief overview – see some more evidence in Chapter 8 – it is clear that respondents need support when they are supposed to in remember their past accurately. We report on the measures developed in the literature in the following section.

2.3 Improving recall in retrospective data collection

Before considering how to improve retrospective data practically, one must think about how memory functions. Relevant for retrieving past events and thus remembering is our so-called autobiographical memory. Theoretical research in cognitive psychology has proposed three main categories of autobiographical memory (see Conway, 1996): *event specific knowledge*, consisting of memories of a particular moment or short period of time such as the event of a car accident or a specific birthday celebration; *general events*, relating to certain periods in one's life, such as a vacation in a specific country or the work for a certain company; finally there are *lifetime periods*, the major periods in one's life, such as childhood, early motherhood, time spent in education, etc.

In our setting of retrospective data collection all of these memories are of interest – we would like to know about lifetime periods, asking generally about a person's working life or the time thereafter. We also are interested in general events, when we ask about a certain job or the time spent in a certain accommodation. Finally, the event specific knowledge plays also a role, when we ask the respondents to recollect memories of health problems or what specifically happened to them during times of persecution.

There are several intuitive ways to improve the recall of an event itself, which have been shown to make a significant difference in recall. Several studies have shown that remembering events is more likely the longer people have time to think about the question. For example, Cannel et al (1977) report that adding meaningless parts to the question increases accuracy at least for educated respondents (with no significant effects for less educated people). Similarly, Loftus et al. (1990) reduce the number of misreports by asking the same question

twice and varying the reference period. Explicitly asking for more effort from the respondent has been shown to increase recall as well (Cannel et al, 1981). However, these approaches are not necessarily suitable for a project like SHARELIFE – interviewers would feel awkward to ask the same question twice for example, and respondents would wonder why a certain question is said to be more important than another one. In addition it is hard to judge whether an event did not happen or the respondent does not remember it, and thus the countermeasures are hard to implement.

If the event in question is remembered, dating accuracy is most important. This is also important for SHARELIFE, as we are interested in precise data. Again, there are some intuitive ways, but also some more elaborate possibilities have been developed in the field. Immediately obvious is the restriction to time periods that are recent – stemming from the results that events that are further back in time are usually less accurately remembered than things that happened more recently. Findings on changing the recall order (i.e. backward, forward or "free recall") are mixed and suggest that the effects of recall order depend on the subject. For example, Jobe et al. (1990), find that free recall provides more accurate results when asking about household health care visits, whereas Loftus et al. (1992) find no significant effect for recall order when asking about visits to health maintenance organizations (HMO). Another option is "bounded recall" – an artificial restriction of the reference period that has been shown to increase accuracy (e.g. Neter and Waksberg, 1964; or Auriat, 1993). Bounded recall can be especially useful in a panel study, where the time between two interviews is naturally defined as the reference period. At the time of the second interview the respondent is reminded of his or her answer given in the first, and then asked to report any changes that happened since. But the reference period can also be restricted to other periods of the respondent's life – for instance to lifetime periods or extended events in the terminology used above. For example, one would ask the respondent to think of his/her childhood and then ask specific questions about it.

Another form of bounded recall is the use of "temporal landmarks", which hinges on the idea that there are certain events in one's life that are outstanding – personal events such as the birth of a child or marriage, but also some public events, such as the assassination of J. F. Kennedy or the win of one's national team in a big sports competition. These events, once they are known, can be used to anchor the respondent's memory and place other personal events relatively to the landmark: the respondent might not know when exactly something happened, but if she knows it was in the year when people landed on the moon, this information can help.

Several studies have tested the use of public landmarks. For example, Loftus and Marburger (1983) use the eruption of Mt. St. Helens in Washington State. Six months after this event, they ask about crime victimization, using either "in the last six months" or "since the eruption of St. Helens" as an entry to the question. They find that the inclusion of landmarks increases the accuracy of the reports.

Besides changing the type or the content of questions, graphical devices have been shown to improve data quality as well. The simplest way here is the use of

timelines, where the respondent is asked to record life events on an axis, and then place other events accordingly around it. This one-dimensional version has been extended to a life grid or calendar with the use of Event History Calendars (EHC), as described by Freedman et al. (1988), Blane (1996), or Belli (1998). The idea is similar to that of timelines, just on multiple dimensions. When going through a questionnaire, life events are recorded into a large grid, where a set of topics such as children, partners, or work are combined with the time dimension, which is usually on the horizontal. The calendar then allows the respondent to see important events of different areas of her life in parallel. Belli (1998) argues that the EHC enhances the respondent's ability to recall, as standard recall mechanisms, which relate back to the types of memories mentioned above, are triggered by this approach. These mechanisms or associations are threefold, all of which are supported by the EHC:

1. Top-down retrieval, meaning that a higher order structure indexing allows moving into lower order structures, or similar, the memory moves from general structures to specific ones. In the framework of the above memory types, one would move from lifetime periods to general events to event specific knowledge. For the implementation in a survey, this suggests to specify large topics first, and then move within these topics to the more specific events.
2. Within a theme, events are ordered along the time dimension and can therefore be recalled sequentially. In designing a questionnaire, one would implement this strategy when asking about recurring events – for example when asking about the different houses a person lived in, one would start with the first (or the last), and then move forward (or backward) along the time dimension.
3. Across themes, recall happens in parallel, meaning that one event can trigger the memory of a different theme because it happened in the same time episode. For survey design, this proposes to somehow visualize to the respondent his or her answers, such that a parallel retrieval is possible – for example, a marriage may coincide with a certain job period.

As Schwarz and Oyserman (2001) suggest, EHCs improve recall by making use of several of the other approaches mentioned above. One clearly is the use of landmarks, as during the process of filling the calendar, the interviewer can always prompt by using previously entered events, for example the birth of children. But also public landmarks can be used very easily, if the interviewer has a possibility to refer to them. The calendar is in principle also open in terms of the order in which it is filled – as long as the questions are asked flexible enough, the topics do not need to be followed in a specific order. An additional feature is that the calendar allows both the respondent and the interviewer to easily cross-check events and correct errors that otherwise would have remained undetected.

The first EHC has been implemented by Freedman et al. (1988) in a pen and paper version. They used this technique in a sample of nine hundred 23 year olds, who were asked in detail about life events that had happened since their 15^{th} birthday. Events were entered on a monthly basis, so overall, 120 entries were possible for each of the possible categories. Since this was done on paper, even

though a lot of categories were included, the scope of the questionnaire accompanying the calendar was rather limited.

Belli et al. (2005) experimented with a life history for the Panel Study of Income Dynamics (PSID). The PSID meant to test the life history approach to see if recall could be improved with this technique compared to a regular q-list interview. The implementation was a computerized instrument done by a telephone interview, so the respondent did not have access to the calendar, but the calendar rather served as an input device for the interviewer. However, the interview itself was very flexible: Instead of asking direct questions, the interviewers were asked to simply give broad questions that would lead to the topics the researchers were interested in. This would guarantee that the respondent truly remembered events. However, this approach is very limited if the researcher is interested in things the respondent cannot be "steered" towards.

The English Longitudinal Study of Ageing (ELSA) implemented a face-to-face interview with an EHC in their study in 2007 (Scholes et al., 2009). As ELSA is very closely related to SHARE, the life history approach serves as a role model for SHARELIFE. The main reason for ELSA to use a life history approach is the same as in SHARELIFE: since all of the respondents are at least 50 years old, the initial conditions for the respondents are not observed. Especially for health and socio-economic status, researchers would be very interested in the respondents past to relate it to the present. In this regard, the life history approach was meant to collect very important unobserved variables.

ELSA covered several areas of retrospective data: health, economics and social networks from early childhood, which were followed then by experiences through adulthood. A lot of development went into the implementation of the instrument, which started out as a very flexible type of interview, similar to the PSID experiment done by Belli et al (2001). However, throughout the course of pre-tests and pilot studies, the study moved from a flexible pen-and-paper combined interview to a more standard way of questioning with a life grid as a supplement to the interview. It is this computerized version of the life history interview that serves as the basis for the SHARELIFE interview described in detail below.

2.4 Overview of Topics

SHARELIFE is meant to analyze the European welfare state by comparing individual decisions across time and countries and connecting these decisions with the institutional surroundings that people faced at the time. This implies a somewhat standardized interview which guarantees the comparability not only within a country but also across the European boundaries. In addition, the face-to-face interviews from the first two waves of SHARE were meant to be continued. These considerations led to the decision not to use a fully flexible approach as taken by Belli et al. (2005) but rather base the life history interview on an approach similar to the English Longitudinal Study of Ageing. This has the additional benefit that ELSA and SHARELIFE are collecting data in a very similar way, allowing researchers to combine the two on numerous dimensions.

There are several different modules to the SHARELIFE interview, which are ordered according to what is usually most important to the respondent and thus remembered most accurately. Although there is a default order, a flexible approach is allowed in the sense that the interviewer can change to any module at any point in time if necessary.

The interview starts in default order with questions about the children, i.e. year of birth of the oldest child, his or her name, gender, etc. Immediately, this information appears in the calendar for both the respondent and the interviewer to see, so that the interviewer has an easy way of linking questions to personal events (in this case, children). The child section is followed by the module about the partner history, which asks about marriages, cohabitating partners as well as about other important relationships. Again, the main information like the start and end of a relationship is displayed on the screen. The places of living are recorded in the following section, where the previously recorded life events prove to be very helpful: interviewers can prompt with that information, e.g. “Did you live there after your second child was born?” or “Were you still with X when you moved?”. This anchoring gives tremendous help to the respondent.

This is followed by a section about the respondent’s living situation when he or she was ten years old. This detailed look at one point in the respondent’s childhood provides useful information about where our respondents come from, as some of these variables prove to be good predictors of socio-economic status later in life. None of the parts of this section appear in the calendar, since the information is too detailed and only concerns this one point in time. This is different for the work section, which follows. This is one of the very detailed sections, where questions are asked about the respondent’s job and retirement history. Not only jobs are covered in detail, but also any periods of not working, be it due to unemployment, maternity or retirement. The work module is followed by a brief section about the use of financial assets during the respondent’s life, where mainly the entry points are of interest.

The next two sections cover health, where the first one is about health status as a child and as an adult, and the second is about health care with a strong focus on the use of preventive medicine in a respondent’s life. These sections allow identifying differences in the actual health and health care use throughout Europe, which are important determinants of the welfare state. The final section of the interview is then covering general life events, where the respondents are asked to identify specific periods of their live, e.g. when they were happier or when they had to endure financial hardship. One important decision here was to include a special section on persecution, even though it may have affected only few of our respondents.

2.5 Summary

While the literature in the second section of this chapter has shown that there are many challenges in collecting retrospective data, the third section has provided good examples that there are multiple ways to rise to the occasion. Indeed, the field has greatly advanced in the past twenty or so years, which now – with the

latest technology available – are possible to implement in actual ongoing studies such as ELSA or SHARELIFE. These studies themselves then test advances on a large scale and thus can contribute to the development of new techniques to the benefit of future studies.

References

Akerlof, George and Janet Yellen. 1985. "Unemployment through the Filter of Memory". *The Quarterly Journal of Economics*, Vol. 100, No 3, pp 747-773.

Auriat, Nadia. 1993. "My Wife Knows Best: A Comparison of Event Dating Accuracy between the Wife, the Husband, the Couple and the Belgium Population Register". *Public Opinion Quarterly*, Vol. 57, pp 165-190.

Belli, Robert. 1998. "The Structure of Autobiographical Memory and the Event History Calendar: Potential Improvements in the Quality of Retrospective Reports in Surveys." *Memory*, Vol. 6 No. 4, pp383-406.

Belli, Robert, William Shay, and Frank Stafford. 2001. "Event History Calendars and Question List Surveys: A Direct Comparison of Interviewing Methods". *Public Opinion Quarterly*, Vol. 65, pp 45-74.

Belli, Robert, Lynette Smith, Patricia Andreski, and Sangeeta Agrawal. 2005. "Methodological Comparisons between CATI Event History Calendar and Standardized Conventional Questionnaire Instruments: Quality of Life Course Retrospective Reports." Working Paper No8, University of Nebraska, Lincoln.

Blane, David, 1996. Collecting retrospective data: development of a reliable method and a pilot study of its use. *Social Science and Medicine*, Vol 42, pp 751-757.

Cannel, Charles, Kent Marquis, and André Laurent. 1977. "A Summary of Studies of Interviewing Methodology". *Vital and Health Statistics*, Series 2, No 26.

Cannel, Charles, Peter Miller, and Lois Oksenberg. 1981. "Research on Interviewing Techniques". In S. Leinhardt (Ed.), *Sociological Methodology*, Vol. 18, pp.37-68. San Francisco: Jossey-Bass.

Freedman, Deborah, Aland Thornton, Donald Camburn, Duane Alwin, and Linda Young-DeMarco. 1988. "The Life History Calendar: A Technique for Collecting Retrospective Data." In C.C. Clogg (Ed.), *Sociological Methodology*,Vol. 18, pp.37-68. San Francisco: Jossey-Bass.

Gaskell, George, Daniel Wright, and Colm O'Muircheartaigh. 2000. "Telescoping of Landmark Events: Implications for Survey Research". *Public Opinion Quarterly*, Vol. 64, pp 77-89.

Jobe, Jared, Andrew White, Catherine Kelley, David Mingay, Marcus Sanchez, and Elizabeth Loftus. 1990. "Recall Strategies and Memory for Health Care Visits". *The Milbank Quarterly*, Vol. 68, No. 2, pp 171-189.

Jobe, Jared, Roger Tourangeu, and Albert Smith. 1993. "Contributions of Survey Research to the Understanding of Memory". *Applied Cognitive Psychology*, Vol. 7, pp 567-584.

Jürges, Hendrik. 2007. "Unemployment, Life Satisfaction, and Retrospective Error". *Journal of the Royal Statistical Society A*, Vol. 170, No. 1, pp 43-61.

Loftus, Elizabeth, and Wesley Marburger. 1983. "Since the Eruption of Mount St. Helens, Has Anyone Beaten You Up? Improving the Accuracy of Retrospective Reports with Landmark Events". *Memory and Cognition*, Vol. 11, pp 114-120.

Loftus, Elizabeth, Mark Klinger, Kyle Smith, and Judith Fiedler. 1990. "A Tale of Two Questions: Benefits of asking More Than One Question." *Public Opinion Quarterly*, Vol. 54, pp 330-345.

Mathiowetz, Nancy and Greg Duncan. 1988. "Out of Work, Out of Mind: Response Errors in Retrospective Reports of Unemployment". *Journal of Business and Economic Statistics*, Vol. 6, No. 2, pp 221-229.

Means, Barbara and Elizabeth Loftus. 1991. "When Personal History Repeats Itself: Decomposing Memories for Recurring Events". *Applied Cognitive Psychology*, Vol. 5, pp 297-318.

Means, Barbara, Arti Nigam, Marlene Zarrow, Elizabeth Loftus, Molla Donaldson. 1989. "Autobiographical Memory for Health Related Events". National Center for Health Statistics, Vital Health Stat, 6(2).

Scholes, Shaun, Jibby Medina, Hayley Cheshire, Kate Cox, Elizabeth Hacker and Carli Lessof, 2009. "Living in the 21st century: older people in England. The 2006 English Longitudinal Study of Ageing Technical Report". National Center for Social Research, NatCen, UK Data Archive Study Number 5050.

Neter, John, and Joseph Waksberg. 1964. "A Study of Response Errors in Expenditures Data from Household Interviews". *Journal of the American Statistical Association*, Vol. 69, No. 305, pp 18-55.

Peters, Elizabeth. 1988. "Retrospective Versus Panel Data in Analyzing Lifecycle Events". *The Journal of Human Resources*, Vol. 23, No. 4, pp 488-513.

Pierret, Charles. 2001. "Event History Data and Survey Recall: An Analysis of the National Longitudinal Survey of Youth 1979 Recall Experiment". *Journal of Human Resources*, Vol. 36, No 3, pp 439-466.

Rubin, D.C. (1996), *Remembering Our Past. Studies in Autobiographical Memory*, Cambridge, Cambridge University Press.

Rubin, D.C. and A.D. Baddeley (1989), "Telescoping is Not Time Compression: A Model of the Dating of Autobiographical Events," *Mem Cognit*, Vol 17, 6, 653-661.

Schwarz, Norbert and Daphna Oyserman. 2001. "Asking About Behavior: Cognition, Communication, and Questionnaire Construction". *American Journal of Evaluation*, Vol. 22, No.2, pp 127-160.

Sudman, S. and N.M. Bradburn (1973), "Effects of Time and Memory Factors on Response in Surveys". *Journal of the American Statistical Society* 68, 344, 805-815.

Tourangeau, Roger , Lance J. Rips and Kenneth Rasinski. 2000. *The Psychology of Survey Response*, Cambridge: Cambridge University Press.

Tulving, Endel. 1983. "Elements of Episodic Memory". Oxford University Press.

Wagenaar, Willem. 1986. "My Memory: A Study of Autobiographical Memory over Six Years". *Cognitive Psychology*, Vol. 18, 225-252.