

Paper prepared for the Census conference "The Use of Event History Calendar Methods in Panel Surveys". Organized by The University of Michigan's Panel Study of Income Dynamics (PSID) and the U.S. Bureau of the Census. December 5-6, 2007, Washington D.C.

**The use of landmark events in EHC-interviews to enhance recall
accuracy**

Wander van der Vaart and Tina Glasner
Wageningen University and Research Centre
The Netherlands

wander.vandervaart@wur.nl

Wageningen University and Research Centre
Department of Social Sciences
Research Methodology Group (Bode 175)
P.O. Box 8130
6700 EW Wageningen
The Netherlands

Abstract

One of the methods that researchers use to aid recall in surveys, consists of employing so called personal landmarks. These are very salient events from peoples lives that aim to facilitate recall of other events and particularly the dates of these events. Personal landmarks are commonly used within calendar methods like timeline procedures and event history calendars. These instruments employ a graphical time frame in which next to the life history information landmarks can be represented. In spite of the fact that sociologists and health scientists apply calendar methods increasingly,

research into their mechanisms and effects on data quality is scarce.

This paper offers some theoretical notions on the application of personal events as landmarks in survey interviews. This leads to the suggestion that the most suitable landmark events probably are: important, domain related, personal events, that are generated by respondents themselves. Data are presented that illustrates what type of landmarks respondents generated spontaneously in a calendar procedure. The data originate from a field experiment on the event history calendar as embedded in the PSID-survey in 1998.

The findings were largely in line with the studies into autobiographical memory and life events. The figures demonstrate that respondents use a great variety of landmark events, that the number and types of landmarks are related to socio-demographic factors and that the retrieval of landmarks shows typical recency and heaping effects. In the present study no clear relationships between landmark usage and recall accuracy were found. The results suggest that an unrestricted use of personal landmarks may not be optimal and that a standardization of the landmark instruction might add to the effectiveness of their aided recall function.

Acknowledgments

We owe a special debt to Robert F. Belli (Gallup Research Center/University of Nebraska) for supporting this project, providing interview transcripts and facilitate data access. The present paper is scheduled as a part of our collaborative work. The study is supported by a National Institute on Aging grant (5R01AG17977-5) awarded to Belli and a Netherlands Organization of Scientific Research grant (400-03-331) awarded to Van der Vaart.

The data partly stem from the public use dataset 'Panel Study of Income Dynamics, Event History Calendar Methods Study'. This dataset is produced and distributed by the University of Michigan with primary funding from the National Science Foundation, the National Institute of Aging, and the National Institute of Child Health and Human Development. Ann Arbor, MI.

We thank Angela Nijland for her share in the coding of interview transcripts.

Introduction

Autobiographical memory, in its widest sense, contains information about the events, which have taken place in a person's life. Among those events, there are "special periods of time" (Shum, 1998) that are more memorable than other periods. These events have been called temporal landmarks, because in a similar way to spatial landmarks, they can be used as reference points, when trying to access other locations in memory. Temporal landmarks can stem from different thematic domains and they can be public as well as personal events. Any event that can be dated fairly accurately and that provides an anchoring point in time can serve as an idiosyncratic memory landmark to a person. In that sense, the definition of "landmark" is functional rather than theoretical/content-based. According to Shum (1998, p. 423), landmark memories serve a passive, as well as an active role. Firstly, these events are stored in memory in a more detailed way than other events; secondly, they help people organize and access other autobiographical memories. Especially the latter aspect has received attention from cognitive psychologists and other researchers who are interested in organizing principles of human memory.

Survey researchers, for instance, may incorporate temporal landmarks into retrospective questionnaires for two purposes:

1. Demarcating the beginning of the reference period;
2. Providing temporal anchors within the reference period.

In past studies, the inclusion of public (Loftus & Marburger, 1983) as well as of personal landmarks (Van der Vaart & Glasner, 2006) has been shown to enhance response quality in surveys. The employment of landmark events forms an important characteristic of so-called calendar methods in social and medical surveys (e.g., Freedman et al., 1988; Sobell et al. 1988). These methods can be used in addition or as an alternative to the standard survey method with chronologically ordered question lists. Calendars provide respondents with a graphical time frame (e.g.: Figure 1) that aims to facilitate access to long-term memory, when answering retrospective questions. Most instruments consist of two-dimensional grids, in which one of the axes

denotes the time dimension while life themes (“work”, “residence”, “education”, “health” etc.) are specified on the other axis. Calendar instruments have been designed for stimulating the respondent to derive content and timing cues from other autobiographical events. They also make it easier for the researcher to discover incompleteness and inconsistencies in retrospective reports.

In recent years the use of calendar techniques such as the Event History Calendar (Belli, 1998) has been growing rapidly in the social and medical sciences (for a review see Glasner & Van der Vaart, in press). The outcomes of evaluation studies indicate that those instruments have beneficial effects on the quality of retrospective data. Also, experimental studies have demonstrated that calendar instruments improved recall accuracy regarding the number of reported events, dates and characteristics of events (Belli et al., 2001, 2004; Van der Vaart, 1996, 2004; Van der Vaart & Glasner, 2007). However, it remains unclear in which way specific design characteristics of calendars contribute to these beneficial effects. In order to shed some light on this issue, our study concentrates on one important characteristic of calendar instruments: the use of landmark events. The aim of the current paper is to gain more knowledge about how to identify and select autobiographical events that can be used as landmarks in calendar instruments.

Few studies have addressed the phenomenon of temporal landmarks in a content-oriented way. When trying to answer the question, which events are the most likely ones to become landmarks in one’s life, researchers have often turned to either short-term studies with student samples and short reference periods, or long-term studies with elderly respondents and very long reference periods. The former type of study has the advantage that events can be dated very precisely within the reference period. On the downside, the sampled population and reference period (usually one year in college) are neither representative of the population as a whole, nor do they reflect the whole range of life events a person can experience. Long-term studies, on the other hand, include a wider spectrum of events, but they usually don’t allow for very precise dating. In our current study, we will try to

combine the advantages of both types of research by having a relatively diverse group of respondents provide landmark events from a two to three year reference period. The landmarks were collected as part of a retrospective survey, in which respondents were asked in the introduction to the questionnaire to report “memorable events” from the reference period.

In this paper we will explore the following questions:

1. How many and which type of landmark events do respondents report in case of unrestricted recall instructions?
2. How are the reported landmarks distributed within the reference period and within calendar years?
3. Do different types of respondents use different types of landmarks?
4. Are numbers and types of reported landmarks related to recall accuracy?

First we will present the main issues from studies into autobiographical memory regarding the retrieval of life events and landmarks. Next we will illustrate what types of landmark events respondents use in survey research by exploring data from a retrospective survey study. The aim of the study is to gain more insight into the aided recall potential of different types of landmarks in survey research.

II. Landmark events and autobiographical memory

In order to understand what type of events might function as landmarks in surveys, we will focus on four main issues from autobiographical memory research. From each issue we will derive basic propositions that may direct our further study into suitable landmark characteristics.

Time-tagged events in autobiographical memory

First, it is important to realize, that only a few events are stored in autobiographical memory with a specific date, often referred to as a “time tag” (e.g. Friedman, 1993). Events that are rehearsed, commemorated or celebrated regularly (like a birthday, one’s own wedding, a nation’s Liberation Day), or events, that occur on “easy to remember” dates, such as 02/02/2002, are most likely to be time-tagged (Janssen, Chessa & Murre, *subm.*). This might imply that time-tags are not necessarily assigned to the memory during encoding but might be added at a later stage, i.e. during rehearsal. Most autobiographical events, however, personal as well as public, do not occur on memorable dates or are salient enough to be time-tagged. As a consequence dating such events involves a certain amount of reconstruction (Friedman, 1993). While time-tagged events are usually dated quite accurately (Burt, 1992), the quality of reconstructed event dates depends on the retrieval strategy. Obviously, reconstructing dates inferring information from untagged, and potentially falsely-dated events, increases the risk that the target event is dated incorrectly. This results in two basic propositions: (1) in order to help respondents reconstruct the date of other events landmarks should be time-tagged events; and (2) landmarks with time-tags probably are often rehearsed events.

Most memories stem from recent periods or from the reminiscence bump

Second, studies that asked people to recall memories from their life, may clarify what type of events in a person’s memory might stand out and how they might function as a landmark. It has generally been found that when older adults are asked to provide autobiographical memories from their lives without restrictions on content or time period, roughly half the reported events

occurred within the most recent year of life (Holmes & Conway, 1999; Rubin et al., 1998); a phenomenon called the “recency effect”.

In addition, there is robust and substantial empirical evidence that people tend to report a relatively great number of events that have occurred in adolescence and early adulthood. This latter phenomenon is referred to as the “reminiscence bump” (Rubin et al., 1998). Holmes & Conway (1999: 22) claim, that “the reminiscence bump is the result of preferential or privileged encoding of experiences relating to two crucial phases of development of the self in adolescence and early adulthood.” In line with this, they found that adults’ recall of public events was best for the first half of the reminiscence bump (10 to 19 years) and autobiographical memories of private experiences are concentrated in the second part (20 to 30 years). Those early and significant experiences are considered to be processed more thoroughly than other events, and form focal points in long-term memory.

In a series of event cueing experiments, Brown and Schopflocher (1998: 310) found similar differences between important and unimportant events. They reported that: “In general, participants retrieved a high percentage of very recent events and few older events when responding to word cues or cueing events elicited by word cues [...]. In contrast, important cueing events and event memories cued by these events were less likely to be very recent, and more likely to come from an earlier period in the participant’s life.” In order to explain this phenomenon Brown and Schopflocher (1998) distinguish between two different forms of autobiographical memory. They put forward that people maintain a pool of recent, mundane events and a pool of older, more important events.

Both phenomena – the recency effect and the reminiscence bump - are important when life events are used as landmark events in interviews. They lead to two propositions: (1) If chosen without restraint by respondents, landmarks might be distributed very unevenly over reference periods. Most of them will stem either from the second and third decade of the respondent’s life or from very recent periods. (2) If (public) landmarks are supplied by the researcher their effects will be quite different for people of different age cohorts. Amongst others, it depends on the respondents’ age and phases of preferential encoding, which public events might serve as landmarks to them.

Lifetime periods as organizational units in autobiographical memory

A third main issue from autobiographical memory concerning the use of landmark events in calendar tools is derived from the fact that most calendar applications are – at least partly- based on Barsalou (1988) and Conway's (1996) models of autobiographical memory. In short, those multi-level models (see also their updates in: Burt, Kemp & Conway, 2003; Conway & Pleydell-Pearce, 2000) state that thematically organized lifetime periods serve as the primary organizational units to autobiographical knowledge. Especially important are those lifetime periods that are central to one's concept of self and its changes across the lifespan, a point that will be further elaborated upon in the next paragraph. Lower in the structure, memories of "general" or "summarized" events, which took place during those lifetime periods can be found, which in turn are grounded in specific experiences. From these models our next proposition can be derived: Landmarks will be most effective if they are focused on the central lifetime periods of the target population.

Narratives as a primary form of organization in autobiographical memory

Finally, a growing number of studies have emphasized narrative-like structures in autobiographical memory, which draw together information about causally or thematically related events (Brown, 2005). Barsalou (1988) observed that summarized events seemed to be chronologically organized and appeared to form a "life narrative". He stated that the relationship in memory between the timelines of different life domains is defined by "a logic of goal attainment".

Brown (2005) and Brown and Schopflocher (1998) provided empirical evidence that people organize autobiographical memory mainly in the form of narratives and narrative-like structures, which the authors call event clusters. Noteworthy events, even minor ones, appeared to be organized in clusters. Clustered events are often causally related, similar in age (i.e., temporally proximate) and referring to the same people, locations and activities (i.e., similar in content). It is important to recognize that clusters do not share just one type of concept (like persons, emotions, et cetera) but may share each of them, and that no single relation dominates in clusters.

If memory is mainly organized in event clusters, this would imply that landmarks should be most effective if they are part of event clusters. This is confirmed by retrieval time studies, which demonstrated that participants had no more difficulty finding related event memories when the cueing events were old than when they were recent. In contrast, retrieval time did increase with remoteness in case of non-cluster mates. Also, it appeared that important cueing events tended to elicit retrieval of cluster mates more often than unimportant cueing events, again regardless of the age of the cueing event (Brown, 2005; Brown & Schopflocher, 1998).

These findings suggest that important events, which are member of a cluster, can be very suitable as landmark cues for other events in the cluster, regardless of the length of the recall period. Also, clustering implies that landmarks should be tailored to the domain, or “narrative”, at hand. This brings us to the next set of propositions: (1) Landmarks should be important events from event clusters. This implies that landmarks should be tailored to the question domain. (2) In search for landmarks one should look for relationships with events; there is no need to look for a specific set of concepts. (3) Landmarks – if embedded in a cluster - are likely to prompt retrieval of events that are temporally proximate and similar in content.

III. Translating the landmark notions to the survey setting

In order to identify specific characteristics of events that might function as a landmark in survey interviews, the issues presented in section II can be complemented by some further survey methods research. Together they point out four characteristics that might enhance the effectiveness of landmarks in survey interviews.

First, findings presented above, indicated that landmarks probably are most effective if they are important events (and not necessarily focused on a specific issue). Important events are more likely to be time-tagged, show stronger cluster effects and are less susceptible to over retrieving recent memories. In line with that Wright and Nunn (2000) found - in a survey study on public events - that importance was the most vital characteristic for maintaining clear autobiographical memories. Mathiowetz and Duncan (1988) distinguished four aspects contributing to the saliency or importance of events in survey settings:

- a. Evoking emotions at the time of occurrence;
- b. Marking a transition point;
- c. The resulting economic or social costs and benefits;
- d. Long term consequences.

Sudman and Bradburn (1983) added the 'unusualness' of an event or state, in the sense of 'occurring rarely'. Although the importance of events contributes to the landmark function and is positively related to recall accuracy, it also may entail negative effects. Importance may enhance telescoping, resulting in displacement of the landmark in time (e.g., Bradburn, Rips & Shevell, 1987). Furthermore, narratives of important events are usually retold more often, which may lead to conscious or unconscious distortions of events. These distortions may in turn result in biased memories (Marsh & Tversky, 2004).

Second, the findings in section II indicate that domain-related landmarks are more useful than unrelated events. This is also true for survey settings. Bradburn et al. (1987) reported that describing a certain time period in personal terms ("your time in high school or college") or in public terms ("during the presidential term of Carter") enhanced dating accuracy for personal and public events respectively.

Third, several studies have shown that personal landmarks are effective recall aids (Means et al., 1988; Jobe & Mingay, 1989) and generally more effective than public landmarks (Baddeley et al., 1978) in eliciting retrieval of other events from autobiographical memory.

This is in line with recent theories of autobiographical memory (Burt, Kemp & Conway, 2003; Conway & Pleydell-Pearce, 2000) which state that most important lifetime periods - those that are central to the self - serve as the primary organizational units of autobiographical knowledge. In addition, memories of personal events do not decrease as strongly as memories of public events and people tend to recall significantly more personal than public events from their life span (Howes & Katz, 1992; Holmes & Conway, 1999). Also memories of personal events – like the number of events and their timing - appear to be less dependent on socio-demographic characteristics of the respondents than memories of public events (Gaskell et al. 2000; Howes & Katz, 1992).

Fourth, it seems that landmarks will be more effective if they are generated by respondents rather than provided by the researcher (Brown, 2005). Howes & Katz (1992) found that participants who were provided with cues by the researcher reported less memories of personal events from their life than those who engaged in spontaneous, unrestricted recall. In the latter condition, recall was less affected by whether participants recalled life events in forward or backward order. In contrast, standardized cues inhibited backward recall.

In the context of survey research, the findings presented above suggest that the most suitable landmark events probably are: important (I), domain related (II), personal events (III), that are generated by respondents themselves (IV).

IV. An empirical exploration of the use of landmarks in a survey interview

In the following sections, we will present results of a study on landmarks used in a retrospective survey, which included a calendar instrument.

Data

The landmarks examined here, originate from a study in which a paper-and-pencil Event History Calendar was used in a telephone survey (Belli, Shay & Stafford, 2001). In order to compare methods, 617 participants in the Panel Study of Income Dynamics (PSID) were randomly assigned to being interviewed either with an EHC, or with a standardized question list. The current study will focus on the 309 respondents in the EHC condition, who had to provide personal landmarks at the beginning of the interview. Data collection took place in May and June 1998, using paper-and-pencil telephone interviewing in both conditions. The reference period covered the years 1996 and 1997. Answers from the previous wave of the panel study were used as control data for establishing the accuracy of respondent's reports regarding events that took place in 1996.

The calendar appeared to have some positive effects on recall accuracy as compared to the question-list. For a number of continuous measures, such as income and weeks away from work, correlations with the control data were higher in the EHC condition than in the question-list condition. The calendar did not differ significantly from the question-list regarding mean signed errors of continuous data, though it did reduce underreporting and enhanced higher reports on some categorical variables. More details about the design of the survey and the effect of the EHC on recall accuracy can be found in Belli et al. (2001).

The Event History Calendar

The calendar instrument used in the study comprised six thematic domains (residence, household, employment, unemployment, and aid entitlements) and covered a reference period of two years. At the start of the telephone interview, immediately after the general introduction of the study, respondents

were asked to list landmark events from the end of 1995 to the date of interview in 1998. Interviewers read out the following instruction (Belli et al., 2001, appendix B):

“Today, it is [May, June] of 1998. We are especially interested in events that had happened from around 2½ years ago up to those that happened around 5-6 months ago. We are especially interested in where and with whom you have lived, your employment and unemployment history, and your income during this time. First though, are there any events in the past few years, from the end of 1995 to the present, that stand out in your mind that you are able to date either exactly or approximately? These might be family, financial, job, social, or health related events.”

Examples of probes that the interviewer could use for eliciting more personal landmarks included “births, deaths, divorces or marriages, a vacation, health-related events such as an accident, major purchases that you have made, a job promotion or pay rise, a residence or job change” (Belli, Shay & Stafford, 2001). Interviewers also had a list of public holidays, such as Christmas and Thanksgiving, which they could offer the respondents as dating cues.

Method and respondents

We analyzed the transcripts of 235 interviews, in which landmarks were collected. For 230 of the respondents, we also had background information on socioeconomic characteristics and the respondent's answers on the main questionnaire. There was a significant difference in age between the 134 female (mean age = 42.2 years) and 96 male respondents (mean age = 47.8 years). Due to oversampling of minority households in the Panel Study of Income Dynamics, 41% of respondents in our sample were black, 54% were white, and 5% reported another race. Even though in the general PSID, different respondents from the same household can be interviewed in subsequent waves, only respondents who had also taken part in the previous wave (1997) were interviewed for the methodological study.

On average, interviewers and respondents spent 3.4 minutes on collecting the memory landmarks, and another 13.9 minutes on the rest of the interview. Data entry after the interview took an average of 29.4 minutes. Information on the operational variables and socio-demographics can be found in Table 9b. The interview time and the time spent on the landmark domain were only weakly related ($r=.20$, $p=.02$). The time needed for data entry is mainly related to interview time ($r=.49$, $p=.00$) and less to the time required for the landmark domain ($r=.18$, $p=.01$). Of the socio-demographic variables, only age appeared to significantly predict the length of the interview, which was shorter for older respondents ($b=-.291$, $t(212)=-4.55$; $p<.001$) and the time it took to collect the landmarks, which was longer for older respondents ($b=.17$; $t(213)=2.67$; $p<.01$). In contrast, data entry time was related to all socio-demographic variables. Taking into account the length of the interview and the total number of reported landmark events, data entry was somewhat lengthier for women ($b =.12$; $t(211)= 1.98$; $p=.05$), white respondents ($b=.11$, $t(211)= -1.76$; $p=.08$), and respondents in the higher income categories ($b=.13$, $t(211)= 2.15$; $p=.03$).

Landmark coding scheme

During the 231 interviews of which we had usable transcripts, respondents generated a total of 535 landmark events. Those events were marked in the printed transcripts and keywords or key phrases were entered into an SPSS database, together with the date of the event. Note that we did not use the original calendars as our data source but transcripts of the first part of the interviews. We only coded events that were mentioned by the respondent, during this initial part of the interview.

Our coding scheme was built in several steps. Based directly on the respondents' description of events, both authors made their own first classification scheme. These schemes aggregated similar individual events into categories. After discussing the differences in both schemes, which were small, the authors formulated 17 event categories. Using this classification, all events were coded independently by three coders: both authors and an assistant. Based on three coders the inter-coder reliability (Hayes &

Krippendorff, 2007; Krippendorff, 2004) was Krippendorff's Alpha = 0.92 for all categories including unclear and not applicable (535 units, nominal measurement level; 231 respondents). Among each pair of coders Alpha ranged from 0.91 to 0.93 and Cramer's V ranged from V= 0,89 to V=0.90. All events that were classified differently by two coders were reconsidered and, after reaching agreement, classified into one of the 17 landmark categories (see Table 1).

Table 1. Basic classification of landmark events.

<i>Code</i>	<i>Label</i>	<i>Content key-words</i>
1	Holiday specific	Holidays (including weekends and trips) with additional information (like destination, activity, persons etc)
2	Holiday generic	Holidays (including weekends and trips) without any specification
3	Health	Health events
4	Births	Births
5	Deaths	Deaths
6	Accidents	Accidents
7	Birthdays own	R's own birthday
8	Birthdays others	Another person's birthday
9	Purchases	Consumer purchases: buying, selling
10	Leisure activities	Leisure events and activities (including sports, music, other hobbies, parties, pets, volunteer work)
11	Public events	National and/or political events, public holidays
12	Housing	Moving and buying, selling, building houses
13	Family and relationships	Family events, marriage, divorce, personal relationships, conflict, raising children
14	Work	Work events, jobs, unemployment, time off
15	Education	Education events, formal schooling
16	Others	Other landmarks
17	Unclear events	Unclear
99	Excluded	

This basic classification scheme was used to construct two more concise classifications. One involved a reduction to 12 categories by merging related categories that contained only a few landmarks. In addition, we developed a more condensed scheme with 7 categories, inspired by similar classification schemes by Holmes and Conway (1999). In most cases we worked with this 7- category scheme (see Table 2).

Table 2. Condensed classification scheme.

1	Holidays
2	Health (including accidents)
3	Family and relationships (including birthdays)
4	Births and deaths
5	Work and education
6	Housing and leisure (incl. purchases)
7	Others (including public events and unclear events)
99	Excluded

V. Results

5.1 The number and types of landmarks reported

The number of landmarks per person

In the 231 telephone interview transcripts we found a total of 535 valid landmarks events. The mean number of landmark events per respondent was therefore 2.32 (sd=1.90), the median number was 2. The number per person ranged from 0 to 11. Table 3 illustrates that respondents reported more events for 1997 than for 1996 and that a large proportion of respondents reported no landmark at all.

Table 3. The number of landmarks reported in total and for the years 1996 and 1997.

	1995-1998	1997	1996
Mean No of events per respondent	2.32	1.18	.76
sd	1.90	1.16	0.92
No landmarks	39 (16.9%)	74 (32.2%)	112 (48.9%)
1 landmark	46 (19.9%)	77 (33.6%)	72 (31.4%)
>= 2 landmarks	146 (63.2%)	78 (34.1%)	45 (19.7%)

The cumulative percentages of all landmark events show that 17% of the respondents reported no landmarks at all, 60% reported at most 2 landmarks,

and 95% of the respondents reported at most 5 landmark events.

The number of landmarks per topic

Not only does the number of reported events differ greatly per respondent (see also the next section), but Table 4 also illustrates that there are large differences in the numbers of reported landmark events between topics.

Table 4. Frequency distribution of the landmark events: 17 categories.

Rank	Landmark category	N events	%
1	Family and relationships	83	15,5
2	Holiday specific	79	14,8
3	Work	70	13,1
4	Births	63	11,8
5	Deaths	54	10,1
6	Health	40	7,5
7	Education	33	6,2
8	Holiday generic	30	5,6
9	Housing	29	5,4
10	Purchases	17	3,2
11	Leisure activities	14	2,6
12	Accidents	7	1,3
13	Public events	6	1,1
14	Birthday own	4	,7
15	Other events	4	,7
16	Birthdays others	2	,4
	Total	535	100,0

It is possible that the respondents' choice of landmark events was somewhat affected by the instructions they received from the interviewer. Except for education (ranked 7th), public events (13th) and birthdays (14/15th), all event categories were mentioned either in the instruction, or in the suggested probes. In the standardized instruction, the interviewer mentions "family (ranked 1st), financial (10th), job (3rd), social (= leisure, ranked 11th), or health related (6th) events", while the probes include "births (4th), deaths (5th), divorces or marriages (included in 'family events'), a vacation (ranks 2nd and 8th), health-related events such as an accident (12th), major purchases that you have made (10th), a job promotion or pay rise (included in

‘work’), a residence (9th) or job change (3rd)”. In line with earlier findings (Howes & Katz, 1992; Holmes & Conway, 1999), respondents reported very few public events, although there is a chance that this can be attributed (in part) to the instructions, which do not mention public events.

When trying to relate instruction and probing behavior to the results, one should take into consideration, that interviewers were instructed to use a flexible interviewing approach. This means that not all respondents received the same probes, and not every interviewer might have been equally inclined to use probing techniques in order to elicit more landmark events from the respondent. A detailed analysis of interviewer behaviors in the original study can be found in Belli et al. (2004).

The landmark categories we found in the data seem to be very much in line with themes from primary life periods that, according to current memory models, are central to autobiographical memory. This is even better represented by Table 5, which summarizes all landmark events into seven categories.

Table 5: Frequency distribution of the landmark events: 7 categories.

Landmark category	N events	%
Births and deaths	117	21.9
Holidays	109	20.4
Work and education	103	19.3
Family and relationships	89	16.6
Housing and leisure	60	11.2
Health	47	8.8
Others	10	1.9
Total	535	100,0

The classification in seven categories bears some resemblance to earlier findings by Holmes and Conway (1999: 25). Their study of the reminiscence bump demonstrated that in case of free recall of important events people often report events that are “closely linked to or a direct part of the respondents own life”. They classified those personal events in five categories: Illness (11%); Relationship (17%), Work and education (29%),

Home and leisure (24%), Births and deaths (20%). Only the ‘holiday’-category is lacking from this classification; a finding that we will later discuss in more detail.

5.2 The distribution of the reported landmark events over time

Frequency distribution by year

The primary reference period of the study covered two years, 1996 and 1997, but respondents were told to recall events from the end of 1995 up until the time of the interview (May/June 1998). This is why not only events from 1996 and 1997 are mentioned. Some respondents even reported events from 1994, which were also coded. In 14 of the 16 cases, where respondents did not specify the year of an event, they indicated that these events occurred “usually every year” or equivalents of such a general “every year” statement. Most of these regular events were holidays. Table 6 shows the distribution of all 533 dated life events over the recall period per type of landmark category.

Table 6. Number of landmarks by year.

Landmark	Every year	1994	1995	1996	1997	1998	Total
1 holidays	10			39	55	4	108
				22.3%	20.3%	7.5%	20.3%
2 health			3	11	24	9	47
			18.8%	6.3%	8.9%	17.0%	8.8%
3 family/relat	4	2	2	29	43	9	89
		50.0%	12.5%	16.6%	15.9%	17.0%	16.7%
4 birth/death		1	6	43	55	12	117
		25.0%	37.5%	24.6%	20.3%	22.6%	22.0%
5 work/educ		1	3	28	60	11	103
		25.0%	18.8%	16.0%	22.1%	20.8%	19.3%
6 housing/leis			2	21	31	6	60
			12.5%	12.0%	11.4%	11.3%	11.3%
7 others/uncl				4	3	2	9
				2.3%	1.1%	3.8%	1.7%
Total	14	4	16	175	271	53	533
		100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

Table 6 demonstrates that the distribution of life events over the recall period (Chi-Square=55.30, df=42, p=0.082) exceeds a chance expectation. A disproportionate number of landmark events date from the more recent part of the reference period, the year 1997. The fact that the number of reported life events decreases relatively rapidly, even within a short reference period, is in line with the literature on long-term autobiographical memory for important events.

However, there are differences between domains (see figure 1). The recency effect is especially pronounced for ‘work and education’ and ‘health’ events, which both show a decline of more than 50% of reported events within one year. For the other main categories, ‘births and deaths’, ‘family/relationships’ and ‘housing’, but also for reports of holidays, memory decay seems to be more gradual.

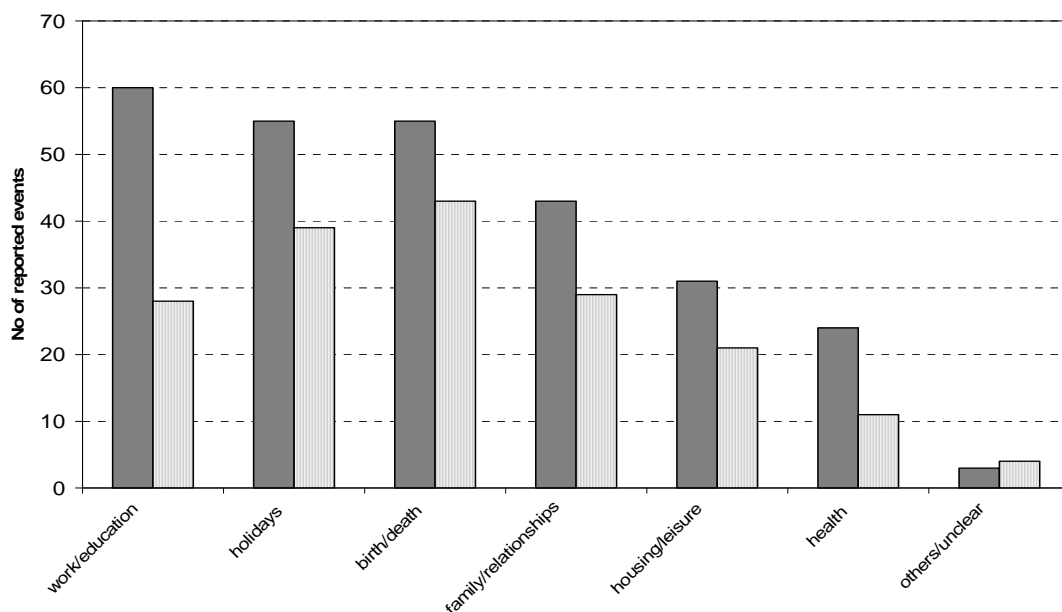


Figure 1: Number of reported landmarks per year (left column: 1997; right: 1996).

These findings (see Table 7) suggest that “matters of life and death”, personally important events, have the highest retention rate in memory. This would mean that over time, while absolute numbers of retrievable events from a specific year are decreasing, the relative importance of these events as memory landmarks will increase.

Table 7. Retention rates (number of events reported for 1996 divided by the number of events reported for 1997) by type of landmark.

Landmark	Retention rate
Birth/death	78%
Holidays	71%
Housing/leisure	68%
Family/relationships	67%
Work/education	47%
Health	46%

While holidays are usually not mentioned as a central category of “important events” in autobiographical memory (see Howes & Conway, 1999), the relatively high “retention rate” of 71% over a period of one year would suggest otherwise. However, a different mechanism might be at work here. In a similar study of memory landmarks, in which we analyzed events from a retrospective reference period of seven years (Van der Vaart & Glasner, 2006), we found that there were two (functionally) different types of holiday landmarks. While reports of “specific” holidays, i.e. holidays for which respondents indicated a destination or travel companion, declined rapidly over the years, reports of “generic” holidays (i.e. “holiday”, without further specification) remained remarkably stable. Also in the present study, though the numbers are not very large, it appeared that the retention rate of ‘generic’ types of holiday events (85%) was higher than for ‘specific holidays (67%). This could be an indicator that, next to being memorable, unique events, (regular) vacations might also be used as time markers within years, similar to school terms, public holidays, or seasonal events.

Frequency distribution by month

The reference period in this study is too short to show ‘reminiscence bumps’, but it’s well known that on a more restricted time scale the dates of retrieved events often show ‘heapings’, for example on specific months within years (Becker & Diop-Sidibe, 2003). In line with that, table 8 demonstrates that the distribution of events over months in our study deviates strongly from a chance distribution (Chi-Square=145,01, df=66, $p < 0.001$). A relatively high number of events were reported to have taken place in either the summer

months or in December. Furthermore, it appears that these ‘heaping effects’ in our data are domain specific. Not surprisingly, more holidays are reported for June, July, and August, than for the rest of the year. Health events, on the other hand, are overrepresented towards the end of the year. Perhaps not surprisingly, most family events were reported for the summer months, and towards the end of the year. The same is true for the category of housing events, leisure, and purchases and, to a lesser degree, also for births and deaths. Comparing the distribution of events over month for 1996 and 1997 shows that there are no large differences between both years (Chi-Square=15.10, df=12, p=0.24).

Table 8. Type of landmark by month (years 1996 and 1997); landmarks, for which no month was indicated, were excluded from the analysis.

Month	Holidays	Health	Family/ relations.	Births/ deaths	Work/ educatio n	Housing/ leisure	Others/ unclear	Total
1	1.1%	6.7%	5.6%	1.0%	3.5%	7.7%	0.0%	3.5%
2	4.5%	3.3%	1.4%	6.3%	2.4%	0.0%	0.0%	3.3%
3	0.0%	10.0%	4.2%	3.1%	5.9%	7.7%	0.0%	4.2%
4	3.4%	10.0%	5.6%	9.4%	7.1%	3.8%	16.7%	6.5%
5	10.1%	6.7%	9.7%	4.2%	18.8%	1.9%	0.0%	9.1%
6	14.6%	6.7%	11.1%	13.5%	16.5%	13.5%	0.0%	13.3%
7	28.1%	6.7%	16.7%	6.3%	4.7%	11.5%	0.0%	12.8%
8	12.4%	6.7%	12.5%	11.5%	11.8%	7.7%	16.7%	11.2%
9	7.9%	6.7%	4.2%	10.4%	11.8%	9.6%	16.7%	8.8%
10	5.6%	6.7%	9.7%	10.4%	2.4%	7.7%	16.7%	7.2%
11	3.4%	13.3%	5.6%	11.5%	5.9%	17.3%	0.0%	8.4%
12	9.0%	16.7%	13.9%	12.5%	9.4%	11.5%	33.3%	11.9%
N	89	30	72	96	85	52	6	430

These findings demonstrate that, if landmarks are self-generated by respondents in an unrestrained way, dates may be biased by domain specific heaping effects. Though not all heapings have to be artifacts (e.g., the heaping of holidays in the summer), the distribution over the months may be very uneven.

Relationships between respondent characteristics and landmark usage

Ranging from 0 to 11, the number of reported landmarks differs greatly per person and the question is whether this is related to socio-demographic factors and whether respondents who report more events from one domain also report more events from other domains. Tables 9a and 9b present an overview. In addition Table 9a illustrates the consequences of using landmarks for interview and data entry time.

There are several weak, but significant, relationships between socio-demographic variables and the number of reported life events (Table 9a). The correlations between socio-demographics and landmarks are similar for 1996 and 1997 reports separately. Gender appeared to be related to the total number of reported landmark events. On average, female respondents (N=133) reported significantly more landmark events than male respondents (N=93) (overall: female 2.59 – male 2.01, $t=2.26$, $p=0.03$). This difference was even more pronounced for events that had taken place in 1996 (female .92 – male .55, $t=3.02$, $p=0.003$). Differences between men and women within the subcategories of events were not significant, except for the holiday category, in which women reported slightly more events.

On average, white respondents reported the most landmarks of all racial groups. While white respondents reported 2.68 events, black respondents reported 1.96 (N=92) and respondents of other races reported 2.17 (N=12) ($F=3.91$, $df=2$, 223, $p=.02$). Taking into account age, gender, and income, the differences between white and non-white respondents as a group are statistically significant for the total number of reported events as well as for events within specific years. With regard to the sub-categories, non-whites reported significantly fewer work and education events.

At first sight, age seems to be negatively related to the total number of reported landmarks. When taking in account race, gender, and income of the respondent, however, this relationship becomes non-significant (the men in our sample were significantly older than the women). The results on specific

domains show a positive relationship between age and the number of reported health events. On all others domains, age is negatively correlated with the number of reported landmarks, even though this correlation is only significant regarding the number of housing events ($p = .03$).

Income is not related to the total number of landmarks reported, but there is a near-significant positive relationship with the number of work and education events ($p = .06$), as well as a near-significant negative relationship with the number of reported births and deaths ($p = .08$).

Table 9a: Multiple regression (standardized betas) of socio-demographic characteristics on the number of reported events for 1996 and 1997: total, per year, and per six event categories (family/relationships & births/deaths combined, 'others' left out).

	Total	1997	1996	Holidays	Health	Family & Births/D.	Work/ed.	Hous.
Age	-.11	-.11	-.12*	-.05	.18**	-.09	-.08	-.17**
Female	.20***	.16**	.24***	.15**	.04	.13*	.07	.04
Non-white	-	-	-	-.12*	-.10	-.13*	-.14**	-.11
Income	.25***	.18**	.23***					
	.01	.05	-.05	.08	.01	-.13*	.14*	-.04
R²	.094	.06	.106	.04	.04	.05	.05	.04
F	5.38	3.4	6.11	1.94	2.35	2.6	2.88	1.86
P	.000	.01	.000	.105	.06	.038	.02	.118

*, **, ***: Significant at .10, .05, and .01 level

Concerning the operational variables, there is a positive correlation between the number of landmarks and the time spent on the landmark domain, also if controlled for socio-demographics (Table 9b). Furthermore, a greater number of landmarks also goes along with longer interviews and with more time spent on data entry, even if controlled for the other variables.

Due to their greater number, the 1997-landmarks show stronger correlations with these operational variables than the 1996-landmarks. Except for Public/other events, all event types are at least slightly positively related to these time measurements. Respondents, who spent more time on the

landmark domain, specifically reported more family events, births and deaths, and work and education events.

Table 9b: Multiple regression of socio-demographic characteristics on operational variables.

	Time spent on landmark domain (Lanlength)	Time spent on rest interview (IWlength)	Data entry minutes (Delength)
IWlength			.35***
Lanlength		.06	.02
Total landmarks	.51***	.36***	.20***
Age	.17***	-.29***	-.10
Female	-.04	-.03	.12**
Non-white	.03	.01	-.11*
Income	-.06	-.02	.12**
R²	.385	.255	.342
F	22.231	12.18	15.73
P	<.001	<.001	<.001

*, **, ***: Significant at .10, .05, and .01 level

Relationships among types of landmarks used

As could be expected, the number of landmarks in each category is correlated with the total number of landmarks, as well as with the number of landmarks for the separate years (Table 10). However, there are few correlations between the number of events as reported in the different categories (over all years). Only the number of birth/death events correlates weakly with the number of family events ($r=.139$, $p=.035$) and there is a slight correlation between the number of work/education events and the number of housing events ($r=.135$, $p=.040$). Apparently, people who report specific types of landmarks are not particularly inclined to report other types of landmarks too. This result suggests that reporting landmarks is not just an indication of a respondent trait, but that the reported events are indeed specific events that stand out in peoples lives.

Table 10: Correlations between the seven types of events reported and the total number of events (all years); N-respondents=213.

	Holidays	Health	Family/ Birthdays	Births/ Deaths	Work/ Education	Housing/ Purchases	Public/ others
N total landmarks	.541(**)	.209(**)	.463(**)	.467(**)	.469(**)	.354(**)	.122

** Correlation is significant at the 0.01 level (2-tailed).

5.3 Relationships between the number and types of landmarks and recall accuracy

The number and types of landmarks and recall accuracy

The event history calendar interviews from which our landmark data were derived, led to somewhat better quality of retrospective reports for eight out of 15 measures as compared to a regular questionnaire (Belli et al, 2001). In the present section we examine whether it is possible to relate the use of landmark events to the beneficial EHC effects. Agreement scores – treated as an indication of ‘recall accuracy’ - could be assessed for the 1996 reports only, by comparing the experimental 1998-reports with the 1997 reports about the preceding year. Some measures, however, reflect a slightly different reference period, which is the period between the interviews held in 1996 and 1997. Although we expect that especially the 1996 landmarks could be related to accuracy measures for retrospectively reported events from 1996, it is not unlikely that the 1997 landmarks also play a role in enhancing data quality. Due to those landmarks, respondents may realize that an event did not take place in 1996 but in 1997, or the other way around. Since the reference period of the PSID-study was 1996-1997 respondents might use landmarks from both years to distinguish between the years. Furthermore, it appeared that a fairly large proportion of respondents did not report any landmark at all, which made it impossible to perform some of the analyses for 1996 and 1997 separately.

Difference scores correlated with landmark frequency

With respect to five measures (having moved, having received child aid, having received food stamps, persons leaving the residence, and weeks out of work due to unemployment) agreement scores were above 90%. In these cases, the remaining 'inaccuracies' were too small to perform analyses on the difference scores. With respect to the remaining 10 measures (Table 11) no correlation was found between the number of landmarks and recall accuracy (in terms of absolute differences) for 6 out of 10 measures. Furthermore, for the remaining 4 measures (regarding the number of weeks away from work due to illness or vacation) the correlations ranged just between $r=.12$ and $r=.15$ and indicated a positive relationship between the number of landmarks and recall error.

Table 11 also presents the relationships between the difference scores and the number of landmarks in specific categories. Since there were great differences between the subcategories work versus education, which formed one category in the 7-category scheme, we present both separately. The category 'others' was excluded due to a very small number of responses in that category. For all 10 measures, there is at least some association between the number of specific types of landmarks (the categories) and absolute recall error. Most associations with absolute error are positive (12 out of 15 times), again illustrating that reporting more landmarks goes along with more recall error.

We've to be very cautious interpreting these few and weak association, but thanks to the use of multiple measurements some patterns become visible. One salient aspect is that the results are very domain specific. Landmarks from the working domain appear to be most central to recall error in this study (related to 8 out of 10 issues), which seems logical since a majority of issues is clearly work related. Also salient is the finding that landmarks from the domains 'Health', 'Family', and 'Education' are related to none of the 10 issues. These outcomes on the types of landmarks apply to all landmark events reported; analyses on the separate years (1996 and 1997) were hampered by too small numbers, but showed a very similar pattern.

The bold types variable in Table 11 – and the same applies to tables 12 and 13 - reflect those measures for which a beneficial effect of the EHC-procedure was established (Belli et al., 2001). Regarding these variables possible landmark effects might be visible more clearly. It appears though, that the landmark do not have specific effects for these variables.

Table 11: Correlations between the number of landmark events reported and absolute recall error. Landmarks in subcategories involve all landmarks reported for 1995-1998.

(Bold variables: effect of EHC on difference scores.)

	N 96+9 7	N 97	N 96	Holi days	Birth /deat h	Wrk/ edu	Wrk	Hous/ leis	N
Persons moved in					-.12 (.08)				220
Number of jobs						.17 (.01)	.15 (.03)	.14 (.04)	226
Income					-.14 (.05)	.18 (.01)	.16 (.02)		205
Wks working							.13 (.06)		213
Wks out of labor							.12 (.08)		213
Wks vacation	.14 (.05)	.12 (.07)				.14 (.04)	.13 (.05)		213
Wks illness self				.12 (.09)			.13 (.07)		213
Illness other	.17 (.01)	-.13 (.06)		.14 (.04)					213
Illness self & other	.13 (.06)		.13 (.07)	.15 (.03)			.12 (.09)		213
Total weeks away (ill & vac)	.14 (.04)	.14 (.04)				.16 (.02)	.17 (.01)		213

Further exploration showed that usually the number of landmarks (total and per category) was only related to recall error if they were also related to the underlying behavior regarding the domains in 1996. Most of the time there's a positive association between the reports on the activities in the life domains and the number of landmarks. All correlations are weak, ranging from $r=.12$ to $r=.31$.

Difference scores by landmark frequency

Table 12 presents the difference scores between the 1998 experimental reports and the 1997 standard of comparison for three categories of respondents, those who had reported no landmarks, 1 or 2 landmarks, and 3 or more landmarks for the reference period 1996-1997. The table presents the same 10 measures as table 11, added by two dichotomous and skewed variables that were not suitable for correlational analysis.

Probably due to the fact that a large proportion of respondents reported no landmarks, the outcomes were much weaker for the separate years. For both 1997 and 1996, significant differences were found for three of the 12 issues (Table 12).

The difference scores complement the results in Table 11 by adding some non-linear relationships between recall error and the number of landmarks. However, there does not seem to be a clear pattern in the outcomes. On the one hand, respondents who reported no landmarks performed worse than other respondents on accuracy measures for 5 indicators: income, number of jobs (signed error only), persons moved into the household, weeks unemployed and whether one had moved residence. On the other hand, on two of these indicators, persons who reported only 1 or 2 landmarks performed better than those who reported 3 or more landmarks. Furthermore, respondents who had reported no landmarks performed best on at least 5 issues (number of jobs (absolute error), weeks worked, weeks out of labor, weeks other ill and weeks total ill).

Again no particular relationship arises between those variables for which a beneficial effect of the EHC-procedure was established and the landmark results.

Based on the results from tables 11 and 12 one can conclude that there is no clear relationship in this study between the number of landmarks reported and the difference scores concerning the experimental scores and the standard of comparison for 1996.

Table 12: Recall error (signed and |absolute|) regarding by the number of landmarks reported for 1996-1997; controlled for age, gender, race and income. (Separate results for 1997 and 1996 only if significant. For dichotomous and skewed variables, two landmark categories were used.) (Bold variables: effect of EHC on difference scores.)

	N landmarks 1996-1997	0	1-2	>= 3	F	p	Df 2/182
1	Persons moved in Persons moved in	.27	.11	.23	2.81	.02	
	<i>N-1997</i>						
	<i>Persons moved in</i> Persons moved in	-.06	.02	.17	2.75	.07	
		.25	.05	.24	3.09	.05	
2	Number of jobs Number of jobs	-.11	-.01	.05	1.90	.08	
		.16	.26	.30	2.23	.04	
3	Income Income	-946	146	481	1.85	.09	
		3151	2879	2925	4.03	.001	
4	Weeks worked Weeks worked	1.81	4.69	2.79	2.35	.03	
5	Weeks out of labor Weeks out of labor	1.51	4.18	2.25	2.54	.03	
	<i>N-1997</i>						
	Weeks out of labor	2.27	5.50	1.35	2.29	0.10	
6	Wks vacation Wks vacation	.73	.70	1.31	1.97	.07	
7	Wks self ill Wks self ill 						
8	Wks other ill Wks other ill	.03	-.10	-.27	5.11	.000	
		.04	.18	.36	6.13	.000	
	<i>N-1997</i>						
	Wks other ill	.06	-.19	-.24	6.86	.001	
	Wks other ill	.07	.22	.33	3.45	.03	
	<i>N-1996</i>						
	Wks other ill	-.06	-.18	-.21	2.46	.09	
	Wks other ill	.11	.32	.30	3.99	.02	
9	Wks ill total Wks ill total 	.39	.39	.92	3.25	.01	
	<i>N-1996</i>						
	Wks ill total	.38	.67	.87	2.53	.08	
	<i>N-1996</i>						
	Wks unemployed	1.19	.40	.49	1.87	.09	
10	Wks away total Wks away total	1.07	.92	1.64	1.84	.10	
	Dichotom /skew variables	0 - 1	>= 2		F	p	df 1/177
11	Wks unemployed Wks unemployed	-1.05	.25		4.17	.04	
12	Moved adress Moved adress	.10	-.02		5.89	.02	
	<i>N-1996</i>	0	>= 1				
	Moved adress	.09	.03		4.41	.04	

Correlations between experimental and standard measures by landmark frequency

The lack of variance in the difference scores is probably an important explanation for the absence of clear landmark effects. Table 13 therefore presents the correlations between the 1998 experimental reports for 1996 and the 1997 standard of comparison. We controlled for age, gender, race and income. It appears that for 8 out of 15 measures the correlations are higher for respondents who reported landmarks from 1996, compared to those who did not report any landmarks. This also applies to the 1997 landmarks for even 10 out of 15 measures. These findings suggest that the overall pattern in the experimental data closer resembles the standard of comparison if respondents did report landmarks. These results were found for those variables for which an EHC-effect was found as well as for variables for which not such an effect was found.

Conclusion

The main finding is that in terms of difference scores no clear relationships can be found between 'recall accuracy' and the number of landmarks reported. However, in terms of associations between the experimental reports and the standards of comparison, respondents who did report landmarks produced higher correlations - in a majority of the cases - than respondents who did not report any landmarks.

Possible the lack of convincing landmark effects is due to the fact that the landmarks were retrieved during the introductory part of the interview and not applied actively during the interview. Furthermore they formed just one component of the whole calendar procedure. In addition, on many measures a large proportion of respondents had difference scores of zero, which means that general dating accuracy was high and little variance was left to be explained. The question remains however, whether the landmarks had any effects on recall accuracy as such. It might be that a greater number of landmark events is just an indication of more activity in people's lives.

Table 13: Correlations between the 1998 experimental reports on 1996 and the 1997 standard of comparison, by the number of landmarks reported for 1996 and 1997; controlled for age, gender, race and income. (Bold variables: effect of EHC on correlations.)

Partial correlations		N96		N97	
		0	>= 1	0	>= 1
1	Persons moved in	.52	.72	.18^a	.83
	N	(95)	(96)	(58)	(133)
2	Persons moved out	.58	.95	.93	.88
	N	(95)	(96)	(58)	(133)
3	Number of jobs	.69	.57	.41	.67
	N	(96)	(99)	(60)	(135)
4	Income	.94	.85	.79	.96
	N	(95)	(97)	(60)	(132)
5	Weeks worked	.77	.64	.64	.74
	N	(94)	(89)	(57)	(126)
6	Weeks out of labor	.74	.67	.68	.74
	N	(94)	(89)	(57)	(126)
7	Wks vacation	.35	.16 ^b	.28 ^c	.26 ^d
	N	(94)	(89)	(57)	(126)
8	Wks self ill	.48	.77	.63	.76
	N	(94)	(89)	(57)	(126)
9	Wks other ill	.58	.69	.69	.68
	N	(94)	(89)	(57)	(126)
10	Wks ill total	.51	.77	.60	.76
	N	(94)	(89)	(57)	(126)
11	Wks unemployed	.75	.28 ^b	-	.83
	N	(94)	(89)	(57)	(126)
12	Wks away total	.48	.53	.55	.52
	N	(94)	(89)	(57)	(126)
13a	Moved reported (i)	.57	.82	.55	.78
	N	(95)	(99)	(59)	(135)
13b	Moved address (i)	.57	.73	.49	.74
	N	(87)	(89)	(53)	(123)
14	Food stamps (i)	.80	.62	.78	.68
	N	(96)	(99)	(60)	(135)
15	Child aid (i)	.64	1.00	-	.81
	N	(96)	(99)	(60)	(135)

- All p < .001, except: a) p = .16, b) p = .01, c) p = .03, d) p = .003
- Spearman correlation coefficient all other: Pearson's correlation coefficient

VI. Conclusion and discussion

Theoretical notions derived from studies on autobiographical memory lead to several suggestions concerning the optimal way of applying life events as landmarks in survey interviews. Notions on time-tagged events, recency and reminiscence effects, the organization of autobiographical memory and the central role of life time periods and narrative-like structures, were combined with survey methods research and transformed into four possibly optimal landmark characteristics. In short the findings suggest that the most suitable landmark events probably are: important, domain related, personal events, that are generated by respondents themselves. Up till present the relative effectiveness of different types of landmarks and how to apply them as recall aid in calendar instruments has not been examined yet

Lessons for DEWS

The usage of landmarks in a survey calendar application

Our empirical exploration of the use of landmarks in a PSID calendar application demonstrated that respondents – in a relatively unrestrained task - use a great variety of landmark events and that there are large differences in the number of landmark events per topic. In our study we inductively came up with seven landmark categories (holidays, health, family/relationships, birth/death, work and education, housing/leisure, others) that had very much in common with the primary life periods that according to current memory models are central to autobiographical memory. Concerning the retrieval of landmark events we found many details that were in line with the studies into autobiographical memory and life events:

- There is a relatively large decline in the number of reported landmark events when the reference period is extend from one to two years;
- Very important topics (birth and death) and generic events (non specific holidays) are relatively stable in time;
- The distribution of landmarks over the twelve months show heaping effects, the location of the heaping depends on the domain at hand;

- Associations (though weak) exist between the number and types of landmarks and respondent characteristics (mainly gender and race, also age);
- In the present study the life events from certain domains are hardly associated with landmarks from other domains (only ‘family’ and ‘birth/death’ events were correlated and also ‘housing’ and ‘work/education’ events).
- The time spent on retrieving landmarks is not related to interview length or data entry length: when landmark have beneficial effects, their cost may be low.

Effects on accuracy

Although we did explore the relationship between the use of landmark events and recall accuracy, we knew that the study wasn’t designed for that aim and that it would be hard to find such associations. To begin with, the landmarks were just one aspect of the whole calendar procedure that we studied. There was no reason to expect that the aided recall effect of a whole EHC-procedure would also be related to one of its parts. In addition, there were large differences between the ways interviewers probed for landmarks during the introduction. Furthermore, the landmarks were hardly used by the interviewers during the core of the interview – which was held by telephone – to actively cue the respondent’s memory. Thus asking for landmarks during the introduction foremost might have functioned as a “warming up” procedure that stimulates recall.

The drawbacks mentioned above, as well as the relatively small number of landmarks per respondents and the low levels of variation in the difference scores, probably explain why the landmarks reports hardly show systematic associations with recall accuracy. However, positive effects of landmarks are suggested by the higher correlations between the 1998 experimental reports and the 1997 standard of comparison for respondents that did report landmarks, in comparison to those who did not report landmarks. It’s not clear though, whether this difference is due to recall effects. The findings that

indicated a negative relationship between landmark reports and accuracy at least present a warning. The application of landmarks may also bias results and thus needs proper tests.

The procedure

The outcomes of both the theoretical and the empirical findings of this study make clear, that an unrestricted or flexible use of landmarks as recall aid in calendar procedures - and in survey interviews in general - is not optimal. Standardization of the landmark instruction probably adds to the effectiveness of the landmark function. Allowing respondents to recall any type of landmark events from any time period opens the procedure to response biases, such as the tendency to recall events from the most recent period or from the reminiscence bump (see also: Howes & Katz, 1992; Rubin et al., 1998), and the heaping of events within or over years. Furthermore, spontaneously generated landmarks lead to differences in the number and types of events used by different socio-economic groups of respondents.

In a standardized instruction the researcher could urge respondent to use the type of landmarks that is expected to be most optimal. According to our conclusions drawn from the literature these are: important, domain related, personal events, that are generated by respondents themselves. The instruction could be regulated at least in two ways. First, in order to reduce recency and heaping effects, respondents could be requested to recall landmark events from specific time periods, like from each season, or (equal numbers) for each half year (Howes & Katz, 1992; Rabbit & Winthorpe, 1988). Second, requiring 'domain related' landmarks means that respondents are asked to recall landmarks from some specified themes. This is in line with the important role of lifetime periods as organizational units in autobiographical memory and the narrative-like structures of knowledge in memory (Brown, 2005; Burt, Kemp & Conway, 2003). This implies that asking respondents to recall landmarks from each specific theme or topic that is central to the data collection, should enhance the recall of more effective landmarks. Presenting loose examples of landmarks and probing in a loose

ways for landmarks might become very steering, as is also suggested by data in the present study. Regulating the instruction might prevent this.

Closing remarks

The answers to our four research questions on landmark usage by respondents in a survey setting - as put forward in the introduction - can be summarized as follows.

- 1) Respondents use many different types of landmark events, though they fit well to the central life domains.
- 2) Most events stem from the most recent year and are unevenly distributed over the months. Both the retention rates and the heapings are domain specific.
- 3) Female and white respondents reported more landmarks than male and non-white respondents. Age, race and income influences the types of landmarks.
- 4) Landmark usage was weakly related to recall accuracy and showed both positive and negative associations with accuracy.

If landmarks are to be employed as aided recall tools, it is important that landmark events are dated accurately and that they do not disturb or bias retrieval processes. From both laboratory studies into autobiographical memory and questionnaire studies it appears that a free recall of life events may not lead to optimal landmarks. In contrast, they suggest that landmarks are probably most effective if they are employed in an interview setting that entail a standardized landmark instruction rather than a flexible or unrestrained procedure. Empirical data in the current study illustrated this need and provided an impression of the applicability of landmark events as recall aids in survey interviews.

References

- Baddeley, A.D., Lewis, V., & Nimmo-Smith, I. (1978). When Did You Last...? In M.M. Gruneberg, P.E. Morris, and R.N. Sykes (eds.), *Practical Aspects of Memory*. New York: Academic Press.
- Barsalou, L. W. (1988). The content and organization of autobiographical memories. In U. Neisser, & E. Winograd (Eds.), *Remembering reconsidered: Ecological and traditional approaches to the study of memory* (pp. 193-243). New York: Cambridge University Press.
- Becker, S., & Diop-Sidibe, N. (2003). Does Use of the Calendar in Surveys Reduce Heaping? *Studies in Family Planning*, 34(2), 127-132.
- Belli, R. F., Lee, E. H., Stafford, F., & Chou, C.-H. (2004). Calendar and question-list survey methods: Association between interviewer behaviors and data quality. *Journal of Official Statistics*, 20(1), 185-218.
- Belli, R. F., Shay, W., & Stafford, F. (2001). Event History Calendars and question list surveys: A direct comparison of interviewing methods. *Public Opinion Quarterly*, 65, 45-74.
- Bradburn, N.M., Rips, L.J. & Shevell, S.K. (1987). Answering autobiographical questions: The impact of memory and inference on surveys. *Science*, 236, 157-161.
- Brown, N.R. (2005). On the Prevalence of Event Clusters in Autobiographical Memory. *Social Cognition*, 23 (1), 35-69.
- Brown, N.R., & Schopflocher, D. (1998). Event Cueing, Event Clusters, and the Temporal Distribution of Autobiographical Memories. *Applied Cognitive Psychology*, 12, 305-319.
- Burt, C. D. B. (1992). Retrieval characteristics of autobiographical memories: Event and date information. *Applied Cognitive Psychology*, 6, 389-404.
- Burt, C. D. B., Kemp, S., & Conway, M. A. (2003). Themes, events, and episodes in autobiographical memory. *Memory & Cognition*, 32, 317-325.
- Conway, M.A. (1996). Autobiographical knowledge and autobiographical memories. In D. Rubin (Ed.), *Remembering our past: Studies in autobiographical memory* (pp. 67-93). Cambridge: Cambridge University Press.
- Conway, M. A., & Pleydell-Pearce, C. W. (2000). The construction of Autobiographical memories in the self-memory system. *Psychological Review*, 107, 261-288.
- Freedman, D., Thornton, A., Camburn, D., Alwin, D., & YoungDeMarcco, L. (1988). The life history calendar: A technique for collecting retrospective data. *Sociological Methodology*, 18, 37-68.
- Friedman, W. J. (1993). Memory for the time of past events. *Psychological Bulletin*, 113, 44-66.
- Gaskell, G.D., Wright, D.B., & O'Muircheartaigh, C.A. (2000). Telescoping of Landmark Events. Implications for Survey Research. *Public Opinion Quarterly*, 64, 77-89.
- Glasner, T.J., & Van der Vaart, W. (in press). Applications of Calendar Instruments in Social Surveys. A Review. *Quality and Quantity*. DOI: 10.1007/s11135-007-9129-8.
- Hayes, A. F., & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures*, 1, 77-89.

- Holmes, A., & Conway, M.A. (1999). Generation Identity and the Reminiscence Bump: Memory for Public and Private Events. *Journal of Adult Development*, 6 (1), 21-34.
- Howes, J.L., & Katz, A.N. (1992). Remote Memory: Recalling Autobiographical and Public Events from across the Lifespan. *Canadian Journal of Psychology*, 46 (1), 92-116.
- Janssen, S. M. J., Chessa, A. G., & Murre, J. M. J. (subm.). Scale effects in memory for time. *Acta Psychologica*.
- Jobe, J.B. and Mingay, D.J. (1989). Cognitive Research Improves Questionnaires. *American Journal of Public Health*, 79, 1053-1055.
- Krippendorff, K. (2004). *Content analysis: An introduction to its methodology*. Thousand Oaks, CA: Sage.
- Means, B., Mingay, D.J., Nigam, A., & Zarrow, M. (1988). A Cognitive Approach to Enhancing Health Survey Reports of Medical Visits. In M.M. Gruneberg, P.E. Morris and R.N. Sykes (eds.), *Practical Aspects of Memory: Current Research and Issues*. Chichester: John Wiley.
- Loftus, E. F., & Marburger, W. (1983). Since the eruption of Mt. St. Helens, did anyone beat you up? Improving the accuracy of retrospective reports with landmark events. *Memory & Cognition*, 2, 114-120.
- Marsh, E.J., & Tversky, B. (2004). Spinning the Stories of our Lives. *Applied Cognitive Psychology*, 18, 491-503.
- Mathiowetz, N.A. & Duncan, G.J. (1988). Out of work, out of mind: Response errors in retrospective reports of unemployment. *Journal of Business & Economic Statistics*, 6 (2), 221-229.
- Rubin, D.C., Rahhal, T.A., & Poon, L.W. (1998). Things Learned in Early Adulthood are Remembered Best. *Memory & Cognition*, 26 (1), 3-19.
- Shum, M.S. (1998). The Role of Temporal Landmarks in Autobiographical Memory Processes. *Psychological Bulletin*, 124 (3), 423-442.
- Shum, M.S., & Rips, L.J. (1999). The Respondent's Confession: Autobiographical Memory in the Context of Surveys. In M.G. Sirken., D.J. Herrmann, S. Schechter, N. Schwarz, J.M. Tanur, and R. Tourangeau (Eds.) (pp. 95-109), *Cognition and Survey Research*. New York: Wiley.
- Sobell, L. C., Sobell, M. B., Leo, G. I., & Cancilla, A. (1988). Reliability of a timeline method: Assessing normal drinkers' reports of recent drinking and a comparative evaluation across several populations. *British Journal of Addiction*, 83, 393-402.
- Sudman, S. and Bradburn, N.M. (1983). *Asking Questions: A practical Guide to Questionnaire Design*. London: Jossey-Bass.
- Van der Vaart, W. (1996). *Inquiring into the past. Data quality of responses to retrospective questions*. Ph.D.-thesis. Amsterdam: VU University.
- Van der Vaart, W. (2004). The time-line as a device to enhance recall in standardized research interviews: A split ballot study. *Journal of Official Statistics*, 20(2), 301-317.
- Van der Vaart, W. & Glasner, T.J. (2006). *Using personal landmarks to enhance recall in survey interviews*. Paper presented at the XVI ISA World Congress of Sociology, Durban.
- Van der Vaart, W., & Glasner, T.J. (2007). Applying a Timeline as a Recall Aid in a Telephone Survey: A Record Check Study. *Applied Cognitive Psychology*, 21 (2), 217-238.

Wright, D.B., & Nunn, J.A. (2000). Similarities within Event Clusters in Autobiographical Memory. *Applied Cognitive Psychology, 14*, 479-489.