# Cognitive processes in Event History Calender interviews: A verbal report analysis

Tina Glasner[1], Wander van der Vaart[2]

[1,2]Wageningen University and Research Centre
Department of Social Sciences
Research Methodology Group
[1]e-mail: tina.glasner@wur.nl
[2]e-mail: wander.vandervaart@wur.nl

## Abstract

For this project, we conducted 37 cognitive interviews. In 29 of these interviews, calendar instruments were used, either as a recall aid in addition to a standardized retrospective questionnaire (n=14), or as a stand-alone data-collection device (n=15). During the interview, which covered a reference period of 20 years, respondents were asked to think aloud while answering retrospective questions and/or filling in the calendar. The interviews were transcribed and we subsequently coded verbal protocols for of 34 interviews in SequenceViewer 4.4 (Dijkstra, 2008). Our coding scheme covered cognitive processes involved in answering retrospective questions about event dates, such as sequencing, top-down retrieval, and the use of landmark event cues. The comparison revealed that even though event dating processes did not differ significantly between conditions, subjects used the calendar to check the quality of their responses and made corrections based on the visual feedback.

Keywords: Event History Calendar, cognitive interviews, retrospective questions

## 1. Introduction

Life History Calendars have become increasingly popular instruments for collecting retrospective data in survey interviews. Various different types of those instruments have been developed in the social and medical sciences. Usually, they consist of two-dimensional grids, which include a time-dimension, several thematic domains, and personal and/or public landmark events, to which the timing of autobiographical episodes can be related. Some calendar instruments are used as recall aids, others as stand-alone data collection devices, in which answers are recorded during the interview and coded by the interviewer at a later point. Calendar instruments have been used with various modes of data collection, including paper-and-pencil interviews (e.g.Becker & Sosa, 1992; Freedman, Thornton, Camburn, Alwin, & Young-DeMarcco, 1988; Goldman, Moreno, & Westoff, 1989), telephone interviews (Belli, Shay, & Stafford, 2001); (Freedman et al., 1988), and computerized personal interviews (Pebley & Sastry, 2004). Traditionally, calendar instruments have been administered by interviewers, often using a relatively flexible interviewing approach (Glasner & Van der Vaart, forthcoming). Due to the complicated structure of most

instruments, few self-completed calendars have been tested (W. Van der Vaart & T. J. Glasner, 2007; Wiebe & Landis, 2000). With the current advance of online survey technology, however, using calendar instruments as automatic visual feedback tools in online questionnaires might become an attractive new option for survey researchers.

So far, a very limited number of studies have evaluated the effectiveness of calendar techniques, and no study has compared different formats. A review of the methodological literature on calendar techniques supports the view that those instruments do have some positive effects on response quality (Glasner & Van der Vaart, forthcoming). When used for paper-and-pencil data recording, they enhance completeness and consistency of retrospective life course data (Becker & Sosa, 1992; Goldman, Moreno, & Westoff, 1989) and improve the timing of autobiographical episodes (Belli, Shay, & Stafford, 2001). Calendar recall aids have also been shown to have positive effects on the timing of purchase events (W. Van der Vaart & T. J. Glasner, 2007).

## 2. Research objective

The use of calendar instruments originates in the research practice. Even though early reports were quite accurate about the sort of effects that the calendar might have on data quality (Balán, Browning, Jelin, & Litzler, 1969), there was a clear lack of theoretical foundations. This situation changed when Belli (1998) linked the Event History Calendar to the structure of autobiographical memory, thereby providing a rationale for its assumed (and partially proven) effectiveness. Building onto the work of Conway (1992) and Barsalou (1988), Belli suggests that calendar instruments aid autobiographical recall by offering the respondent sequential, top-down, and parallel cues. While this assumption is often cited and seems to be generally accepted by researchers who use calendars tools, it has never been confirmed in a systematic way. Also, no comparisons have been made between the cognitive processes involved in answering retrospective questions in regular survey questionnaires, and in interviews in which calendar instruments are used. With this study, we intend to start filling the gap between cognitive rationale and research practice by taking a closer look at Event History Calendars using a cognitive interviewing approach.

Belli, Lee, Stafford, & Chou (2004) made a first attempt at identifying cognitive processes in real-life EHC interviews by coding interviewer and respondent verbal behaviours in 216 EHC interviews and 197 conventional question-list interviews. The authors found that respondents in the calendar condition were more likely to spontaneously use sequential and parallel dating strategies than respondents in the question-list condition. Interviewer probes also differed between conditions with interviewers using more parallel and public holiday cues in the calendar condition. Nonetheless, in total numbers, far more top-down cues than EHC specific cues were used in both conditions.

While providing valuable observations of interviewer-respondent interactions and behaviors, behavior coding is limited to describing observable, outward behaviors. This means that insights into the respondent's thought processes very much depend on the likelihood of the individual respondents spontaneously voicing their thoughts in a "natural interview setting". Behavior coding studies might therefore miss important

clues regarding the retrieval strategies used by respondents in response to calendar recall aids. Cognitive interviewing, another tool for pre-testing survey questions, can often be a valuable addition to behavior coding. Unlike behavior coding, cognitive interviews are more of a "laboratory" approach, and they are usually analyzed in a qualitative way. Their advantage over behavior coding is that respondents are asked explicitly for information regarding cognitive processes. Our study adds to Belli et al's research by shifting the focus from interviewer-respondent interactions to cognitive processes, and by limiting interviewer influences on retrieval strategies.

The aim of this study is to investigate which retrieval strategies survey respondents use when answering retrospective questions in interviews with either Event History Calendars or regular question lists. We also pay attention to the different components of the Event History calendar, time dimension, landmark events and domain grid and their effect on cognitive processes.

We will start with an overview of the EHC's theoretical rationale, followed by a description of the cognitive interviewing method and the coding scheme we used in our study. Subsequently, we present the results of the interviews and connect them to the theoretical rationale.

## 3. Theoretical background: Autobiographical memory and the Event History Calendar

Event history calendars have been theoretically linked to hierarchical models of autobiographical memory, which are based on the assumption that memories of events are embedded in a context of other autobiographical experiences. Conway's (1996) model of autobiographical memory states that events are stored on several interconnected levels of abstraction. On the top of the hierarchy, there are very long-term extended events or "lifetime periods", such as working for a certain employer or living in a certain city. Within those lifetime periods, different "themes" (e.g. work, relationships etc) can be distinguished. One step lower in the hierarchy, there are the memories of summarized events, or "general events", which took place during those lifetime periods. These events vary in specificity. The structure includes memory of more general events such as reading a lot of books, as well as more specific memories, such as working on a certain project, going on holiday, or meeting one's partner for the first time. Memories of these general events are anchored in the "phenomenological record", the memory structure in which very specific phenomenological experiences are stored. Those experiences include sensory (e.g. 'it was cold that day') as well as emotional information (e.g. 'I was feeling disappointed').

### 3.1. Time in autobiographical memory

In autobiographical memory only a few events are time-tagged, i.e. stored in memory with a specific date. This form of representing temporal information is also called a *calendar representation* (Huttenlocher, Hedges, & Bradburn, 1990). Events that are stored with a time-tag are usually retrieved quite accurately (Burt, 1992). With regard to public events, time-tagging is evident for events that have been rehearsed so often that they are actually referred to by their date, such as "9/11" (terrorist attack on the

World Trade Center in New York, on September 11th, 2001) or "3/11" (terrorist attacks on passenger trains in Madrid, on March 3rd 2004), although day and month of those events are probably remembered more reliably than the year in which they took place. Other events might be remembered well, because they are commemorated every year. In terms of personal memories, only some very salient events are time-tagged. Once again, events that are rehearsed and commemorated regularly are probably most likely to be stored with a specific date. These events would include the birth of a child, which is likely to be celebrated on the same day every subsequent year, one's own wedding (anniversary), or high school graduation (reunions).

### 3.2. Reconstructing dates and durations in survey interviews

Many of the events respondents have to report in survey interviews are not time-tagged, which means that their dates have to be reconstructed during retrieval (Friedman, 1993). One method of reconstructing the dates of autobiographical memories is to connect them to other personal or public events in autobiographical memory. The temporal framework used for dating the events is not provided by a calendar, but rather by an *event sequence* (Huttenlocher et al., 1990).

Firstly one can connect the target event to thematically related preceding and/or succeeding events. In the context of Conway's model of autobiographical memory, this means for example, that a general event like "working on project X" took place *before* working on the current project (project Y), but *after* the summer holiday. Retrieving information in this way, is sometimes called "sequencing", since the event is recalled as part of an event sequence (Belli, 1998). Apart from providing us with (somewhat global) temporal information, sequencing strategies also help us to contextualize events, and report them as a narrative, hereby reducing the risk of omitting events.

Secondly, it is possible that the event, that has to be retrieved, happened in close temporal proximity to a public or personal landmark event, which does contain a time-tag. An example from the personal domain would be 'getting one's driver's license two weeks after one's 18th birthday'. Next to personal events mentioned by the respondent (such as weddings, vacations, or birth of children), many instruments also offer public event cues. For relatively short reference periods, those cues can include public holidays or institutional events, such as the beginning of a school term; if the reference period is longer, there will be only few very memorable public events which can be used as landmarks. Even though public events are usually not dated very accurately (Janssen, Chessa, & Murre, 2006), and respondents seem to be reluctant to spontaneously use them as temporal anchoring points (W. Van der Vaart & T. Glasner, 2007), they have two significant advantages above personal landmark events:

a) They can be pre-specified in the instrument as standardized cues;
b) Their dates can be verified by the researcher.

Thirdly, the links between events in a sequence can be logical or causal. If one of the events that took place in the respondent's life enabled, facilitated or prevented other events, the sequence can be used as a retrieval cue.

Next to sequential dating strategies in which information from the same thematic domain is used, the date of an event or episode can also be inferred from other

domains within the structural model of autobiographical memory. This means, that if an event falls into a clearly defined lifetime period or other extended event, the start and end dates of that period form the temporal boundaries within which the other event is located. For example, if a person started her university education in September 1987, the event "buying a car during my second year at university", must be located between September 1988 and September 1989. In relation to Conway's model, Belli (1998) calls this way of inferring dates "top-down retrieval". He also suggests that this type of retrieval might be especially likely to occur, when respondents have to retrieve isolated events, which they cannot place in an event sequence.

Fourthly, the temporal boundaries within which the event is placed do not necessarily need to be derived from more abstract events. It is also possible that timing cues are obtained from other thematic structures within the same lifetime period or extended event (e.g. "When I called off my engagement, I lived in X and had just started working as a teacher"). This process is called *parallel retrieval* (Belli, 1998).

Finally, the timing of autobiographical events can also be based on temporal information inferred from the phenomenological record. If one remembers, for instance, that an event like "moving to Amsterdam" took place while it was freezing outside, one can usually assume that it was winter.

The event dating strategies mentioned above are not mutually exclusive. As a matter of fact, in most cases respondents will use multiple, if not all, strategies for reconstructing dates and durations of past events. Furthermore, use of multiple dating strategies does not necessarily mean that the respondent actually needs more than a single dating cue; it does not even mean that the event is not time-tagged. If the respondent is motivated to date an event as accurately as possible, he or she might retrieve dating cues even after providing a sufficient answer, in order to 'double-check' the accuracy of their statement.

Conventional survey questionnaires are usually structured in a way that prompts top-down and sequencing techniques, but does not explicitly stimulate respondents to make parallel or logic-based inferences from other thematic domains. Calendar instruments, on the other hand, offer a variety of recall cues, which help the respondent retrieve and date autobiographical events and episodes (Belli, 1998). Firstly, all timeline and calendar applications contain landmark events, which serve as temporal bounding cues. Secondly, when calendar instruments are used for recording episodes, they allow for sequencing within life domains. The respondent could be asked for example, to name in forward or in backward order, all employers he or she has worked for, and date those employment episodes (for an example see Engel, Keifer, & Zahm, 2001). Finally, the visual properties of the calendar make it possible for respondents and interviewers to link episodes across life domains, thereby encouraging parallel and top-down retrieval. When a calendar is used next to a standardized questionnaire, the respondent's choice of retrieval technique is likely to be influenced by the question format (most often sequential), but should also contain some of the other elements.

In the current study, we examined retrieval processes reported by respondents in three different conditions. In the first condition, respondents filled out a standardized retrospective questionnaire. In the second condition, they received the same questionnaire combined with a calendar recall aid, which was filled in by an interviewer. Respondents in the third condition completed the calendar themselves and did not receive a standardized question list.

## 4. Method

In order to gain more insights into the cognitive processes involved in reconstructing autobiographical event sequences using a calendar instrument, we collected verbal protocols from survey respondents, using a cognitive interviewing approach. Cognitive interviews are usually focused on questionnaire design properties and aimed at detecting potential cognitive problems that respondents might experience during the interview. These problems can occur at different stages of the question answering process, and might for example concern task comprehension, the retrieval of information, or the use of response alternatives. Often, a non-representative group of subjects, chosen either from the specific target group of the survey, or from those who are expected to be at special risk of experiencing cognitive difficulties (e.g. elderly or impaired respondents), will take part in the pre-test. Their essential task is to provide verbal reports of their thought processes while answering the survey questions tested in the study. There are various ways of collecting verbal report data; the two most commonly used methods in cognitive interviews are “think-aloud” interviewing, in which the interviewer asks the subjects to verbalize their thoughts while thinking about their response, and verbal probing, in which concrete cognitive issues (such as the interpretation of question wording etc) are addressed.

The goal of our current study was somewhat different from that of a regular pre-test. Rather than identify problems, we wanted to describe the cognitive processes respondents engaged in during the retrieval of autobiographical events in survey interviews. We used a think-aloud technique, with relatively restrictive interviewer instructions concerning the use of verbal probes. Permitted probes were essentially restricted to non-directive general encouragements, such as “Would you please think-aloud while you answer the question?” or “Can you tell me what your thoughts were right now?”. The interviews were conducted in November 2006 at Vrije Universiteit Amsterdam by the two authors and by a research student. All interviews were recorded on camera as well as on DVR, and verbal protocols were transcribed by assistants after the interview.

### 4.1. Sample

In cooperation with another Dutch research institute, 37 respondents from the Amsterdam area were recruited from a large online panel. All our subjects took part in online surveys on a regular basis and were contacted after they had filled in the standard electronic questionnaire. The average age of our sample was 54 years (with a minimum of 33 and a maximum of 80 years).

| Age | Number of subjects |
|---|---|
| <40 | 8 |
| 40-49 | 7 |
| 50-59 | 12 |
| >59 | 10 |
| | 37 |

Table 1: Subjects by age

We interviewed 20 female (average age: 53.1 years) and 17 male panel members (average age: 55 years). Higher educational categories were overrepresented in our sample, which nonetheless included a number of subjects with little formal education. Unfortunately we were not able to match respondents across conditions for any other characteristics than gender.

### 4.2. Developing test versions of questionnaire and calendars

The current study was conducted as a pre-test of a calendar method to be used in an online panel survey. To the awareness of both authors, it is the first time for a calendar instrument to be included in web-based data collection. There are several characteristics of this method of data collection that need to be taken into account, when designing the new calendar instrument. First of all, the final electronic questionnaire will be self-completed by respondents, which means that instructions must be clear, and the instrument needs to be easy to use, as there is no interviewer present who can offer clarification. Secondly, the calendar will be part of a computerized questionnaire, which means that the amount of information that can be clearly represented in the calendar is limited by the size of a standard computer screen. Ideally, there should be no need for respondents to scroll in order to see different parts of the memory aid. Finally, the sample will include respondents of all ages and educational levels. This means that the calendar instrument must be accessible to respondents who have little or no experience with computer use and the Internet. The diversity in the sample could have another effect on the design of the calendar. If the research is aimed at gathering information about the respondents' whole life course, the size of the calendar will depend on the respondent's age.

Taking these requirements into account, it seemed unfeasible to design a calendar instrument that is used in the same way as many interviewer administered calendar questionnaires, in which events and episodes are recorded and edited by clicking on the visual display of the calendar instrument. The resulting instrument will therefore not be used as a stand-alone questionnaire, but as a computerized visual recall aid in combination with a more traditional standardized questionnaire. Respondents' answers will be used to build up a calendar, which will provide visual feedback.

### 4.3. Conditions

For the pre-test, our sample was split into three conditions: the first group of respondents filled in a retrospective questionnaire, with a reference period of almost twenty years, starting from January 1st, 1987 and ending at the time of the interview, in November 2006. The second group of respondents filled in the same questionnaire. In this condition, the interviewer also recorded the respondent's answers in a paper-and-pencil calendar instrument, which the respondent could use as a visual recall aid during the interview. Respondents in the third condition did not fill in a questionnaire, but recorded their answers in an open format calendar instrument. Condition two and three were further split into two subgroups each, in which different calendar formats were tested. The difference between those formats lay in the type of landmark events that the calendar offered the respondent. Respondents in the calendar conditions of our study (2a, 2b, 3a, and 3b) were asked to provide any number of memorable events from the twenty-year reference period, that they could date with some confidence, in other words "time-tagged" events. Instructions were deliberately kept open with regard to what kinds of events the respondents were allowed to report, stating only that the event should be "memorable". However, if respondents only reported events from a limited period of time, the interviewer could ask them to report events from other periods, again without specifying the type of event. The subjects in subgroups 2b and 3b received an additional timeline with pre-specified public events from the same period.

Table 2 shows the number of respondents in each experimental group:

| **Condition** | **Number of respondents** |
|---|---|
| 1. Questionnaire only | **8** |
| 2. Questionnaire plus calendar | **14** |
| a) Personal landmarks only | (5) |
| b) Public and personal landmarks | (9) |
| 3. Calendar only | **15** |
| a) Personal landmarks only | (7) |
| b) Public and personal landmarks | (8) |
| Total | **37** |

Table 2: Respondents by condition

### 4.4. Choosing public landmark events for the calendar

In calendar studies, public event cues are usually not chosen in a systematic way. Our intention was to list one memorable event for each of the years in our reference period (1987-2006). We tried to select potential landmark events as objectively as possible, but had to realize that there seemed to be no scientific studies in the Netherlands that met our needs. Fortunately, we were able to locate a popular study conducted by an online research agency for a Dutch public broadcaster, in which the "most memorable" news events of the years 1980 to 2004 were determined, using an online poll among "more than 5000 [sic]" Dutch internet users (https://n40.noties.nl/peil.nl/; accessed 11/08/2006).

In the interest of clarity, we only used events from the 'top 25' section of the list, which meant that the public events timeline was left blank for the years 1988, 1993,

1994, 1998, and 1999. Even though there was no top-ranked event for 1987 either, we decided to include a lower-ranked event (the INF treaty between United States and Soviet Union; ranked 37th) in order to provide respondents with a bounding cue for the beginning of the reference period. If multiple events were mentioned for a particular year, we selected the highest-ranking one; except for the years 1991 and 2004, for which we included two high-ranked events each. As the poll had been conducted in 2004, there were no events listed for the two most recent years of our calendar's reference period. For 2005, we added the death of Pope John Paul II, because of the media attention the event received. By (perceived) lack of memorable events in 2006, we decided to include the football World Cup, that had taken place in neighboring Germany that year, which would give us the opportunity to observe respondents' reactions to sports events as public landmark cues.

### 4.5. Development of the coding scheme

As the main objective of our study was to identify cognitive processes involved in answering retrospective questions, our coding scheme was developed accordingly, with the great majority of codes focusing on the retrieval of dates of autobiographical events. The first category of codes denoted time- or age-tagged events. In the second category, we specified codes that were derived from the model of autobiographical memory discussed in the previous section. These codes were assigned to respondents' verbal reports of retrieval processes, in which memories of other events were used as dating cues. Our third category of codes was not directly event-related. This category included the "causal link" code, which could be assigned to retrieval strategies in combination with codes from category 1 or 2, and codes, which could be assigned to general dating strategies such as estimates of elapsed time or event durations, either alone or in combination with other codes.

Next to the retrieval codes, we specified a number of process codes, in order to identify potential interviewer influences on the respondent's thought process as well as on the respondent's adherence to the think-aloud protocol.

Table 4.1. Time-tagged events

| Code | Description |
|---|---|
| Time tag | a) Respondent indicates that he/she ‘just knows’ the date of an event<br>b) Respondent mentions specific date (dd/mm/yy) |
| Age tag | Respondent uses his or her age as a dating cue (i.e. “I was 23 years old when I graduated, so that must have been in 1993”) |
| Recently retrieved | The event was retrieved recently, e.g. in the process of updating one’s curriculum vitae, and that is why the respondent still knows the date at the time of the interview |
| Copy | The event was mentioned as a personal landmark, and the date is copied from the landmark domain during the interview |

Table 4.2. Cues related to the hierarchical structure of autobiographical memory

| Code | Description |
|---|---|
| Sequencing | Respondent explicitly refers to preceding or subsequent episodes as part of the dating strategy |
| Top-down | Higher-level information (especially lifetime episodes) is used to date a specific event<br>a) Shorter episode is mentioned as part of a longer episode (e.g. several jobs with the same company)<br>b) A specific date is located within a longer episode |
| Personal landmark | a) Respondents uses personal event from personal landmark domain as temporal reference point for dating other events<br>b) Respondent uses personal event which was not mentioned before as temporal reference point |
| Public landmark | a) Reference to public events pre-specified in the public event domain<br>b) Reference to public events mentioned in the personal/memorable event domain<br>c) Reference to public events not mentioned before |
| Same domain event | Respondent refers to event from the same thematic domain in order to date another event or transition |
| Parallel domain event | Respondent uses cue from one of the other substantive domains, i.e. housing, work, education, general health and hospitalizations |
| Phenomenological record | All temporal cues taken from lower level phenomenological information including:<br>a) Physical circumstances<br>b) Clothing<br>c) Weather<br>d) Activities<br>e) Feelings and moods<br>f) Other phenomenological cues |

Table 4.3. Additional codes

| Code | Description |
|---|---|
| Calculation | Respondent mentions retrieval process that includes durations, temporal distances between events, or elapsed time since the event took place |
| Causal links | Respondent refers to causal link between events (e.g. within or across thematic domains) |
| Estimation | Respondent indicates that he/she is not sure about the date but is willing to provide a rough estimate ("I am not quite sure, but it must have been somewhere in 1992. Don't know why, just an estimate.") |
| Season | Respondent mentions only year and season, but either does not specify a month at all, or indicates that the month is arbitrarily chosen from the months within a season (e.g. "I know it was spring, so let's just say April") |
| Recurrent events | Event occurs regularly (e.g. every year) |

Table 4.4. Process codes

| Code | Description |
|---|---|
| Correction | Respondent corrects her/his earlier answer |
| Visual | Respondent explicitly uses the calendar's visual features as a recall aid (e.g. "Mmh, there seems to be a gap… that cannot be right.") |
| Simplify | Respondent consciously 'simplifies' trajectories by ignoring or skipping shorter, more complicated/diverse episodes |
| Don't know | Respondent indicates that he/she does not remember the date of a specified event at all |
| Unclear | Vital parts of the think-aloud and/or transcription are missing, so that the coder cannot determine the retrieval strategy |
| Undefined | Strategy is clear but does not fit any of the other categories |

The correction and visual codes are not mutually exclusive.

Table 4.5. Interviewer behaviour

| Code | Description |
|---|---|
| Think-aloud | Interviewer encourages the respondent to verbalize their thoughts (e.g. "Please think aloud", or "Please don't forget to tell me what you are thinking while you answer the next question") |
| Probe | Interviewer uses retrospective probe (e.g. "How did you remember that date?")<br>If multiple general probing questions were asked after each other, they were coded as separate probes |
| Undesirable interviewer behaviour | Interviewer violates survey protocol by suggesting answers, using suggestive probes, digressing from the subject of the interview et cetera |

Except for the general process codes mentioned in table 5 (interviewer probes, et cetera), only the respondent's utterances were coded. We coded all utterances connected to the retrieval process, even if those utterances were made after the subject had already produced a date for the event. If the subject reiterated a previous comment, the reiteration was not coded again, unless new components were added.

### 4.6. Method of analysis

All 37 interviews were transcribed by assistants and the transcripts were checked by the authors. At this stage, three of the interviews had to be discarded, because the interviewer had not appropriately encouraged the respondent to verbalize their thoughts while retrieving the dates of events, resulting in incomplete or unclear verbal protocols. The remaining 34 interviews were coded in SequenceViewer 4.4 (Dijkstra, 2008) a program best known for its application in behavior coding studies. Recently added text analysis features allowed us to assign text keys to our verbal protocols, and compare their occurrence across conditions. The program also enabled us to correlate text keys with a number of socio-demographic background variables, such as age, gender, and the number of life events and transitions that respondents reported during the interview.

## 5. Results

Before interpreting the results of our study, it should be mentioned again, that our samples in the three categories were very small, which made it almost impossible to find significant differences between conditions. Also, unfortunately for our comparison, the number of reported events differed between conditions. We assume that this is partly due to the fact that we did not have sufficient background information on our respondents to match the conditions with regard to age group. There was a significant negative correlation between age and the number of reported transitions for the twenty-year reference period (where transition is defined as any event that had to be dated during the main interview, mostly start and end dates of autobiographical episodes, or transitions between consecutive episodes). As respondents in the question-list condition were on average almost seven years older than respondents in the two calendar conditions (60.5 versus 52.2 and 55.5 years respectively), this is likely to have influenced the number of reported transitions per condition, which in turn might have affected the number of reported retrieval processes.

For the three main categories, respondents reported the following mean number of transitions for the twenty-year period:

| Question-list (n=8) | EHC plus question-list (n=13) | EHC (n=13) | Total |
|---|---|---|---|
| 7.25 | 11.54 | 13.23 | 11.18 |

Table 5.1. Number of reported transitions per condition

Based on the assumption that the number of reported thought processes depends on the number of events that one has to retrieve, we weighted the results by the number

of transitions, i.e. assigned the weight factor *number of transitions within group/mean number of transitions across groups.*

Unfortunately for our analysis, we did not have sufficient control data to check if the lower numbers of reported transitions in the questionnaire-only category were due to actual differences in experienced events between groups of respondents or caused by the omission of events.

### 5.1. Overview of retrieval processes reported in the three main conditions

In the following, we present an overview of the cognitive processes reported by respondents in the three main conditions, grouped into the categories of our coding scheme.

Table 5.1.: Time-tagged events (weighted by number of transitions)

| | Question-list (n=8) | EHC plus question-list (n=13) | EHC (n=13) |
|---|---|---|---|
| Timetag | 1.63 (2.51) | 2.69 (2.61) | 2.31 (1.95) |
| Agetag | 0.63 (0.96) | 0.46 (0.45) | 0.46 (0.39) |
| Recently retrieved | 0.25 (.39) | 0.46 (.45) | 0.08 (.07) |
| Copied from landmarks | 0.00 | 0.46 (.45) | 0.85 (.72) |
| Total | 2.50 (3.86) | 4.07 (3.96) | 3.69 (3.13) |

Despite being quite substantial in the raw data, the differences in the number of reported time-tagged events between conditions are rather small when weighted by the number of transitions in each condition. Roughly one third of the events and transitions reported in our study were time-tagged. Subjects who reported more events/transitions, also reported more time-tagged events ($p<.01$), but not significantly more age-tagged events. This could be an effect of age, as older respondents tended to report fewer transitions as well as fewer time-tagged events than younger respondents. These results are in line with the expectation that the respondent's ability to retrieve the dates of time-tagged events from autobiographical memory will depend more on individual factors, such as age or cognitive style, than on the method of data collection.

Table 5.2.: Cues related to the hierarchical structure of autobiographical memory (weighted by number of transitions)

| | Question-list (n=8) | EHC plus question-list (n=13) | EHC (n=13) |
|---|---|---|---|
| Sequencing | 1.63 (2.51) | 3.69 (3.58) | 3.54 (2.99) |
| Topdown | 1.13 (1.73) | 1.54 (1.49) | .85 (0.72) |
| Personal landmark | 2.88 (4.43) | 2.93 (2.83) | 3.15 (2.67) |
| Public landmark | .50 (0.77) | 1.77 (1.71) | 1.0 (0.85) |
| Same domain event | 1.38 (2.12) | 3.62 (3.50) | 1.77 (1.50) |
| Parallel domain event | 1.88 (2.89) | 4.15 (4.02) | 2.77 (2.34) |
| Phenomenological record | 2.50 (3.86) | 2.46 (2.38) | 2.23 (1.89) |
| Total | 11.9 (18.31) | 20.16 (19.51) | 15.13 (12.96) |

Again, the weighted data show little difference in the total numbers of reported retrieval processes between the question lists with and without the calendar recall aid. In the next sections (§5.2. and 5.3), we will discuss these findings in more detail and make a distinction between processes that are related to landmark events and those supposed to be elicited by the calendar's domain grid.

Unlike the first two categories of our coding scheme, the "additional" codes we specified (see table 5.3) do not fall into a shared category from a theoretical point of view. The category comprises basic retrieval and date estimation strategies as well as causal links.

Table 5.3.: Additional codes

| | Question-list (n=8) | EHC plus question-list (n=13) | EHC (n=13) |
|---|---|---|---|
| Calculation strategies | 2.75 (4.24) | 4.38 (4.25) | 2.85 (2.41) |
| Causal links | 1.75 (2.70) | 2.15 (2.09) | 2.15 (1.82) |
| Estimation strategies | 1.38 (2.12) | 2.62 (2.53) | 0.85 (0.72) |
| Season-based estimates | 2.38 (3.66) | 2.31 (2.24) | 1.15 (0.98) |
| Recurring events | 0.13 (0.19) | 0.31 (0.30) | 0.46 (0.39) |
| Total | 8.38 (12.91) | 11.77 (11.40) | 7.46 (6.31) |

Overall, it is quite noticeable that the number of reported cognitive processes per transition (the weighted numbers) was smaller in the EHC-only condition than in the other two conditions. Subjects in that condition reported many transitions but relatively few cognitive processes, calendar-related or otherwise. This points toward the possibility that, the think-aloud technique might have been more difficult to apply in this condition, resulting in lower numbers of verbal reports, without necessarily affecting the quality or quantity of retrieval strategies. An analysis of the process codes reveals that interviewers used significantly more probes in the first two conditions, in which data collection was structured by a standardized questionnaire, than in the third condition, in which data was collected with a relatively "open" calendar instrument (4.9 versus 1.9 probes, $p<.05$). Due to these difficulties with implementing the think aloud technique in the third condition we will focus comparisons of cognitive processes in calendar and question list interviews on the first two conditions: standardized retrospective questionnaires without a recall aid and the same questionnaire with a calendar recall aid.

## 5.2. Landmark events

Respondents in the two calendar conditions (n = 28) reported a total of 417 time-tagged life events, varying from two to 71 events over the twenty-year reference period of the interview. 67 of those 417 events were not mentioned in response to the original request to write down landmarks, but were added by the respondents at a later stage of the interview. Contrary to our findings from earlier studies of landmark events (Van der Vaart & Glasner, 2007), where we found a strong negative influence of temporal distance on the number of reported events, we noticed that the distribution of reported events over time was remarkably stable in the current study, despite the long reference period. There was a significant positive relationship between year of the event and the sequence in which events were reported ($r=.407$; $p<.001$), which

indicates that respondents tended to report older events first, often proceeding in forward chronological order. This correlational analysis only included events reported in response to the initial landmark question, not those that were added later during the interview. Respondents under the age of 50 reported almost the same number of events for the first and the second half of the reference period (9.93 versus 9.86), while older respondents tended to report slightly more events from the second half (5.2) than from the first half (4.1) of that period.

Rather surprisingly, we did not find any evidence that subjects in the calendar conditions used more retrieval strategies based on dating cues from personal landmark events than subjects in the question-list condition. When the data was weighted by number of transitions, the picture was even reversed: subjects who didn't use a visual recall aid were more inclined to report using personal landmark cues than subjects who used the calendar ($p < .05$). Also, women used significantly more personal landmark cues than men (3.47 versus 1.27; $p < .01$). Furthermore, and perhaps even more importantly, there was no correlation between the number of reported landmark events in the two calendar conditions and the number of times that respondents reported using personal landmark-based retrieval techniques.

Respondents in the two calendar conditions reported using slightly more public event dating strategies than respondents in the question-list conditions. In the weighted data, the difference between condition 1 and 2 was significant ($p < .01$). A more detailed analysis of the public landmark related retrieval processes reveals that although many respondents commented on the public landmarks in our calendar, most of them did not use them as retrieval cues for dates of other events. Out of a total of 39 public landmark based retrieval strategies, only four were based on the public landmarks specified in the calendar. More popular public event cues were related to recurring events in institutional calendars such as school years, semesters et cetera (mentioned 14 times in the verbal protocols), or public holidays and festivals, most notably Christmas and New Year's Eve (mentioned 15 times). Apart from those categories, six respondents derived retrieval cues from changes in legislation or policy, which had directly affected their personal situation.

### 5.3. Retrieval cues from the domain grid

Next to the landmark events and the time dimension, the domain grid is a central component of calendar instruments. There, events can be recorded and related to each other temporally and causally. According to the theoretical rationale of the Event History Calendar, the domain grid offers survey respondents sequential, parallel, and logic based memory cues for dating autobiographical events and provides them with visual cues, thereby enabling them to detect and, if necessary, remove gaps and inconsistencies in their account. As can be seen in tables 5.2 and 5.3, the discrepancies in the number of sequential, parallel, and logic based cues that respondents used in the different conditions are quite large in the raw data, but become insignificant when take into account the number of transitions. Of course, one should not forget that this is an exploratory analysis, with a very modest sample size, resulting in a lack of statistical power.

When we look at the number of times that our subjects used the calendar's visual features to either date episodes or to correct earlier statements, a somewhat different picture emerges (see table 5.4).

Table 5.4.: Visual feedback

| | Question-list (n=8) | EHC plus question-list (n=13) |
|---|---|---|
| Visual cues | Not applicable | 2.15 (2.09) |
| Corrections | 0.25 (0.39) | 1.85 (1.79) |
| Simplification | 0.0 (0.0) | 0.38 (.37) |

These results indicate that, on average, subjects in condition 2 used the visual feedback provided by the calendar almost four times per interview, either for detecting gaps in the timeline, or for correcting earlier answers in the light of new information. Subjects in the question-list corrected themselves significantly less often. However, they were also less likely to admit to simplifying complex event sequences than subjects who used a calendar (although total numbers are low in this category).

## 6. Conclusion

Despite its obvious limitations in terms of statistical power and randomization between conditions, our study has provided us with some useful insights into the working of Event History Calendars. The results of our comparison indicate that cognitive processes involved in answering retrospective questions might be remarkably similar for regular question-list interviews and interviews, in which calendar recall aids are used. Overall, subjects tended to report multiple thought processes for retrieving the dates of individual events (> 3 in conditions 1 and 2), and about one third of the reported events was time-tagged. Contrary to our expectations, subjects in the calendar condition did not seem to use significantly more cues based on the hierarchical structure of autobiographical memory than respondents who filled in a regular question list. Also, notwithstanding the fact that many subjects did draw on personal event cues when reconstructing the dates of other events, we did not find any evidence that the public and personal landmarks in the calendar had any effect on the retrieval strategies they used. This could mean that it might not be necessary to specify landmark events in a calendar instrument at all. The personal landmarks we recorded before the interview in the calendar conditions might have been accessible as memory cues during the interview regardless of the type of questionnaire that was used. Moreover, even though the subjects in our study noticed the pre-specified public event cues, they all but ignored them when trying to date events during the interview. If public event cues were used at all, those events tended to be either recurrent (such as public holidays) or to have had an impact on the subject's personal situation (such as changes in labor law). All in all, the results of our analysis point towards the possibility that the use of public event cues could be quite idiosyncratic, which would make it difficult to find effective cues for standardized calendar instruments.

However, unexpected as they might be, these results do not necessarily mean that Event History Calendars will not have a positive effect on the accuracy of survey data. Earlier studies have shown, that the instrument tends to improve the completeness of answers and event timing. From our analysis of verbal protocols, we

can conclude that subjects in the calendar condition did use the visual features of the recall aid to check and if necessary correct their answers. In an actual interview setting, the interviewer can also use the calendar in order to detect inconsistencies and gaps in the respondents account and improve data quality by asking for clarification.

From our point of view, further studies on the effectiveness of calendar instruments should focus on the question which components of the calendar (domain grid, personal and public landmark events) are actually necessary to achieve these improvements in the quality of retrospective reports. If it is possible to maintain the positive effects of the Event History Calendar while simplifying the method, these insights can be used in order to develop more cost-effective calendar instruments.

## Acknowledgment

This study was supported by the Netherlands Organization of Scientific Research (grant number 400-03-331).

## References

Balán, J., Browning, H. L., Jelin, E., Litzler, L. (1969) A computerized approach to the processing and analysis of life histories obtained in sample surveys, *Behavioral Science, 14*(2), 105-120

Barsalou, L. W. (1988) The content and organization of autobiographical memories, in: *Remembering reconsidered: Ecological and traditional approaches to the study of memory*, Neisser, U. & Winograd, E. (Eds.), Cambridge University Press, 193-243

Becker, S., Sosa, D. (1992) An experiment using a month-by-month calendar in a family planning survey in Costa Rica, *Studies in Family Planning, 23*(6), 386-391

Belli, R. F. (1998) The structure of autobiographical memory and the Event History Calendar: Potential improvements in the quality of retrospective reports in surveys, *Memory, 6*(4), 383-406

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

Burt, C. D. B. (1992) Retrieval characteristics of autobiographical memories: Event and date information, *Applied Cognitive Psychology, 6*, 389-404

Conway, M. A. (1992) A structural model of autobiographical memory, in: *Theoretical perspectives on autobiographical memory*, Conway, M. A., Rubin, D., Spinnler, H. & Wagenaar, W. A. (Eds.), Kluwer Academic Publishers

Conway, M.A. (1996) Autobiographical knowledge and autobiographical memories, in: *Remembering our past: Studies in autobiographical memory*, Rubin, D. (Ed.), Cambridge University Press, 67-93

Dijkstra, W. (2008) *Sequence Viewer*, version 4.4. URL: http://www.sequenceviewer.nl

Engel, L., Keifer, M., Zahm, S. (2001) Comparison of a traditional questionnaire with an icon/calendar-based questionnaire to assess occupational history, *American Journal of Industrial Medicine, 40*, 502-511

Freedman, D., Thornton, A., Camburn, D., Alwin, D., Young-DeMarcco, 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

Glasner, T.J., Van der Vaart, W. (forthcoming) Applications of Calendar Instruments in Social Surveys. A Review, *Quality and Quantity*

Goldman, N., Moreno, L., Westoff, C. F. (1989), Collection of survey data on contraception: an evaluation of an experiment in Peru, *Studies in Family Planning, 20*(3), 147-157

Huttenlocher, J., Hedges, L., Bradburn, N. (1990) Reports of elapsed time: bounding and rounding processes in estimation, *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*(2), 196-213

Janssen, S., Chessa, A., Murre, J. (2006), Memory for time: How people date events. *Memory & Cognition, 34*(1), 138-147

Pebley, A. R., Sastry, N. (2004), *The Los Angeles Family and Neighborhood Survey: Field Interviewer Manual*: RAND

Van der Vaart, W., Glasner, T. (2007), *The use of landmark events in EHC-interviews to enhance recall accuracy*, US Census Bureau Conference "The Use of Event History Calendar Methods in Panel Surveys". Washington, D.C.

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

Wiebe, E., Landis, K. (2000) *Evaluation of an electronic event history calendar*, Paper presented at the 55th Annual Conference of the American Association for Public Opinion Research & World Association for Public Opinion Research, Portland, Oregon