

Applications of calendar instruments in social surveys: a review

Tina Glasner · Wander van der Vaart

Received: 1 July 2006 / Accepted: 1 January 2007 / Published online: 5 October 2007
© Springer Science + Business Media B.V. 2007

Abstract Retrospective reports in survey interviews and questionnaires are subject to many types of recall error, which affect completeness, consistency, and dating accuracy. Concerns about this problem have led to the development of so-called calendar instruments, or timeline techniques. These aided recall procedures have been designed to help respondents gain better access to long-term memory by providing a graphical time frame in which life history information can be represented. In order to obtain more insights into the potential benefits of calendar methodology, this paper presents a review of the application of calendar instruments, their design characteristics and effects on data quality. Calendar techniques are currently used in a variety of fields, including life course research, epidemiology and family planning studies. Despite the growing interest in these new methods, their application often lacks sufficient theoretical foundation and little attention has been paid to their effectiveness. Several recent studies however, have demonstrated that in comparison to more traditional survey methods, calendar techniques can improve some aspects of data quality. While calendar instruments have been shown to be potentially beneficial to retrospective data quality, there is an apparent need for methodological research that generates more systematic knowledge about their application in social surveys.

Keywords Data collection · Aided recall techniques · Life course research · Retrospective data · Calendar methods · Timeline · Questionnaire design · Memory

1 Introduction

Retrospective reports in survey interviews and questionnaires are subject to many types of recall error, which may affect their completeness, consistency and dating accuracy (Schwarz and Sudman 1994; Scott and Alwin 1998; Van der Vaart 1996; Van der Vaart et al. 1995).

T. Glasner (✉) · W. van der Vaart
Department of Social Research Methodology, Faculty of Social Sciences, VU University Amsterdam,
De Boelelaan 1081, 1081 HV, Amsterdam, The Netherlands
e-mail: tj.glasner@fsw.vu.nl

In the social and the medical sciences, where many studies focus on the reconstruction of life histories, concerns about this problem have led to the development of so-called calendar instruments, or timeline techniques (Freedman et al. 1988; Sobell et al. 1988). These data collection procedures offer an alternative to regular survey questionnaires, which usually consist of lists of chronologically ordered standardized questions, organized in thematic blocks.

Calendar and timeline methods are aimed at helping respondents gain better access to long-term memory by providing them with a graphical time frame (for an example see the Appendix) in which life history information can be represented (Van der Vaart 2004). This stimulates the respondent to relate, visually and/or mentally, the timing among several kinds of events. Inconsistencies in reports are more easily discovered and one event may prompt the recall of another. Additionally, detailed sequences of events are easier to record since they can be marked graphically in the time frame. In recent years the application of these techniques in social research has been growing rapidly. This fact is illustrated by their integration into large-scale, longitudinal social surveys such as the German Life History Study (Brückner and Mayer 1998); the Panel Study of Income Dynamics in the USA (Belli et al. 2001); and the Process of Social Integration of Young Adults in the Netherlands (Van der Vaart 1996). There are some small (historical) differences between the concepts ‘calendar’ and ‘timeline’, as will be pointed out later, but here we will use the word ‘calendar’ for both methods.

Even though the general assumption is that calendar methods improve data quality, there has been little methodological research into their effectiveness. Given the fact that using these aided recall procedures tends to increase operational costs, more methodological insights are needed. In the past two decades, several authors (e.g., Freedman et al. 1988) have described the method in detail, most of the time focusing on the specific type of calendar, which they used in their own study. Still, little is known about the effects of calendar instruments on data quality, measured in terms of completeness and consistency of the data and the occurrence of dating errors. Only recently, a number of experimental studies have been conducted, the results of which indicate that calendars can indeed be beneficial to data quality.

This article presents an overview of calendar instruments currently used in different fields of research and their effects on the accuracy of retrospective reports. Firstly, we will provide an overview of the most important areas of application and present the instrument’s rationale. Secondly, design features and the suitability of the instrument for different modes of data collection will be described. Thirdly, there will be a detailed discussion of the effects of several types of calendar instruments on data quality. Fourthly, we will present the consequences for operational costs and a summary of interviewer and respondent evaluations of those instruments. Finally, we will draw conclusions from our review and offer suggestions for further research into specific features of calendar procedures.

2 Calendar instruments: applications and rationale

2.1 Overview: different names, similar instruments

Most applications of calendar techniques can be found in life course research in sociology (e.g., Giele and Elder 1998) and in health behaviour and treatment studies (e.g., Turner et al. 1992). Both disciplines focus on the retrospective reconstruction of (series or sequences of) behaviour and events during an extended time period and share similar concerns about

recall bias. In life course studies, respondents are typically asked to recall their life history in terms of employment, education, relationships, residences, financial behaviour, et cetera. Health studies usually focus on more specific topics like risk behaviour (e.g., alcohol and drug use, sexual behaviour) and medical events (e.g., cancer diagnosis, clinical treatments, hospitalisation), although there are examples of epidemiological studies, which also include domains like employment and living conditions (e.g., Engel et al. 2001b).

Reviewing the literature on calendar techniques used in social and medical research demonstrates that neither the instrument nor the terminology has been standardized so far. In the social sciences different names for calendar instruments include Life History Calendar (Freedman et al. 1988), Timeline (Van der Vaart 1996), Life History Matrix and Time Axes (Brückner and Mayer 1998), Event History Calendar (Belli 1998), Life Events Calendar (Hoppin et al. 1998), Illustrated Life History (Balán et al. 1969) and Month-by-month calendar (Becker and Sosa 1992). In the medical sciences simple calendar instruments, which measure only one behavioural domain during a relatively short reference period are usually called 'timelines' (Sobell et al. 1988). The reason for the division between (medical) 'timelines' and 'calendars' is foremost a historical one. Most medical timeline instruments are simplified calendars, measuring only one behavioural domain during a relatively short period.

In principle, all calendar or timeline instruments are based on the same idea, which is to enhance autobiographical recall by providing the respondent with event cues. In the following, the term 'calendar instrument' will be used to refer to all of these instruments. Although different versions of calendar instruments have been developed relatively independently from each other in different fields of research, they share three important characteristics:

- (a) The instrument includes a graphical display of the time dimension. Usually, the reference period is divided into smaller time units, such as years, months or days. The size of those time units mainly depends on the length of the reference period.
- (b) The graphical display encompasses one or more thematic axes, representing the domains concerning which the data is collected.
- (c) The respondent is provided with temporal bounding cues, such as public or personal landmark events.

The number of domains and the length of the reference period can vary to a great extent, but most calendar instruments comprise multiple life domains and reference periods longer than one year. The landmark domain may contain personal landmark events or public event cues (e.g., Hoppin et al. 1998), as well as or a combination of the two.

During the fieldwork calendar instruments can be used either as a separate memory aid or as a data collection device. The latter usually applies to the complicated event history calendars which are applied by the interviewer to administrate the data as well as to aid recall by cross checking answers, providing cues and probing for information (Brückner and Mayer 1998). As an alternative, (uncomplicated) calendars can be applied as a separate aided recall device in addition to a questionnaire (Van der Vaart 1996, 2004).

Calendar instruments are currently used in a variety of fields, including life course research (Axinn et al. 1999; Caspi et al. 1996; Reimer 2004; Smith and Thomas 2003; Van der Vaart 1996), epidemiology (Colt et al. 2001; Engel et al. 2001b; Hoppin et al. 1998; Wingo et al. 1988), family planning studies (Becker and Diop-Sidibe 2003; Becker and Sosa 1992; Goldman et al. 1989; Rosenberg et al. 1983), health behaviour (Wiebe and Landis 2000), sexual risk taking (Martyn and Martin 2003) and domestic violence (Yoshihama et al. 2005). This wide range of topics is also reflected in a great diversity of studied populations.

2.2 Rationale of the calendar methodology

Calendar instruments were originally designed to meet practical research needs, which means that in the early days relatively little attention was paid to their theoretical foundations. This situation changed with the advance of the CASM movement, which focuses on the relationship between ‘cognition’ and ‘survey methods’ (see [Tourangeau et al. 2000](#)). CASM studies demonstrated that retrospective data suffer from recall error like omission, underreporting, dating errors, and biased representations, all aspects that are central to calendar techniques ([Schwarz and Sudman 1994](#); [Van der Vaart 1996](#)). The most direct effort to relate the effective mechanisms of calendar techniques to memory processes was made by [Belli \(1998\)](#). Based on hierarchical models of autobiographical memory ([Conway 1992](#)), Belli argued that the Event History Calendar encourages respondents to place events into a temporal context by relating them in several ways to other (parallel or superordinate) events and episodes.

Firstly, the landmark events used in calendar applications serve as temporal anchoring points, or *bounding cues*. Bounding cues were originally developed in order to demarcate the beginning of the reference period in survey research, but can also serve as more general temporal anchor points within the reference period (e.g., “I bought that car the week before Christmas” or “We moved to New York 2 months after my husband graduated from college”).

Secondly, when calendar instruments are used for recording episodes, they allow for so-called *sequencing* within life domains, which means that an event or episode is recalled as part of an event sequence ([Belli 1998](#)). 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 et al. 2001b](#)). Regarding sequences, standardized interviewer probes can essentially be restricted to “What did you do before/after [episode]?” Even though sequencing strategies can be used with other question list formats, the advantage of the calendar over a regular survey questionnaire lies in the fact that the event sequence can be displayed graphically. Apart from providing respondents with (somewhat global) temporal information, sequencing strategies also help them to contextualize events, and report them as a narrative, hereby reducing the risk of omitting events.

Finally, the visual properties of the calendar make it possible for respondents and interviewers to link episodes across life domains, thereby encouraging *top-down* and *parallel* retrieval. Top-down retrieval occurs, when the respondent tries dating an event *within* the context of a more extended episode or “life-time period” (e.g., “I completed the internship during my junior year in college”). The name top-down retrieval implies that recall cues are retrieved from superordinate memory structures. Parallel retrieval occurs when cues are obtained from *other* (not superordinate) 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”). Especially the option of applying parallel crosschecks is a distinguishing feature of calendar instruments. Example questions could be: “Did the incident occur while you were living in X?” or “Did you start working for Y before or after you finished college?” et cetera. Compared to sequential probing, parallel probing probably requires a more flexible interviewing approach, as the interviewer will have to choose the most suitable parallel domain(s) to refer to.

Some first results by [Belli et al. \(2004\)](#) support the idea that calendar techniques are related to these memory processes. Using verbal behavioural coding they found that the event history calendar enhanced ‘sequential’ and ‘parallel’ retrieval strategies by respondents as intended.

Moreover, positive associations were found between these recall strategies and data quality. These findings are promising and show that more experimental studies on the underlying cognitive mechanism of calendar techniques are warranted.

3 Design features of the calendar and modes of data collection

As can be expected from an instrument that is used in so many different fields, calendar techniques come in many different versions. Since one of the purposes of this review is to identify the effects of several design features on the quality of the retrospective reports, we will first describe the layout of calendars that have been used in social research. After that their application with different modes of data collections will be described.

3.1 Reference periods, number of domains and the representation of landmark events

As calendars are often used in life course studies, some applications encompass many years of a respondent's life history and can range from the year the respondent was born to the time of the interview (Hoppin et al. 1998). However, calendar instruments are also used in surveys with far shorter reference periods. Belli et al. (2001), for example, included a calendar in a biennial social survey, in which respondents answered retrospective questions about the two-year period preceding the interview.

While some calendar instruments measure only one life domain (e.g., Engel et al. 2001a) most calendars used in the social sciences include multiple thematic domains. Those domains are presented as parallel timelines, forming a grid, in which one of the axes (usually the horizontal one) denotes the time dimension while on the other axis themes ("work", "residence", "education", "health" etc.) are assigned to each of the timelines. Additionally, landmark events from the reference period can be written down in the calendar.

Calendar instruments (the so-called 'timelines') used in the medical field, especially in research on addictive behaviors, are often less elaborate. They usually include only one behavioral domain, such as alcohol consumption (e.g., Sobell and Sobell 2003) or gambling (Weinstock et al. 2004). Reference periods are short, and the time-units are usually 24-h intervals, in order to measure daily fluctuations in behavior. As a general rule, calendars with a shorter time-range allow for more detailed recording of events, since the temporal axis can be split into shorter time units (e.g., months instead of years).

One distinguishing feature of calendar instruments is the timeline on which landmark events from the reference period can be noted down. In some studies these events are recorded in the timeline of a domain; this especially occurs when information about only one life domain is collected (Engel et al. 2001b; Hoppin et al. 1998; Rosenberg et al. 1983). However, most studies include a separate timeline to write down landmark events (Belli et al. 2001; Van der Vaart 2004). In order to make the calendar instrument more attractive and more easily understandable for populations with limited literacy, some researchers used icons and toy figures (Engel et al. 2001a), or adhesive pictures (Hoppin et al. 1998) instead of written cues.

To our knowledge, the relative effectiveness of different types of landmarks in calendar instruments has not been examined yet. In one study (Yoshihama et al. 2005) respondents were asked which landmark cues they thought were helpful, in this case in recalling partner violence. Respondents mentioned cues from their relationship history and self-generated landmarks ("significant life events") as the most helpful ones.

3.2 Experiences with different modes of data collection

As will be illustrated below, some calendar instruments are completed by the interviewer, while others are filled in by the respondents themselves. The traditional life history calendar (LHC) is an interviewer-administered paper and pencil questionnaire (Balán et al. 1969; Freedman et al. 1988; Wingo et al. 1988). However, it has been used in personal interviews as well as in CATI situations. Freedman and her colleagues (1988) administered almost half of their LHC-interviews via the telephone, in which case the calendar was not visible to the respondent. It could only help the interviewer to detect inconsistencies in the respondent's reported employment history, and prompt the respondent with personalized cues in order to resolve these inconsistencies. Rather surprisingly, the authors conclude, that "the two modes produced almost the same degree of consistency" (p. 65) between the collected retrospective reports and the control data. Unfortunately, the authors do not elaborate on these results. Given their earlier assumption that the respondent uses the calendar as a visual recall aid, one would expect the instrument to be less beneficial to data quality during a telephone interview.

Next to these paper-and-pencil calendar instruments, computerized versions of calendar methods have been evaluated (Belli 2000; Wiebe and Landis 2000). Similar to paper-and-pencil calendars, the computerized instruments are used in personal interviews as well as with CATI applications. The Panel Study of Income Dynamics (PSID), for instance, includes a computerized Event History Calendar (EHC) together with a CATI questionnaire (Belli et al. 2001). A version of this computerized EHC is used during personal interviews of the Los Angeles Family and Neighborhood Survey (LAFANS) (Pebley and Sastry 2004). The instrument is used for collecting data about a relatively short reference period of 2 years previous to the survey.

Instead of using interviewer-administered calendar instruments, some questionnaires include calendars, which are completed by the respondents themselves. These calendars can be part of either paper-and-pencil questionnaires (Martyn and Martin 2003), or they can be used as visual recall aids during personal (Van der Vaart 2004) or telephone interviews (Van der Vaart and Glasner 2005). In the latter study, the calendar instrument was mailed to respondents prior to the interview. In the advance letter and the accompanying written instruction respondents were asked to fill out the simplified life history calendar, which spanned a period of over 7 years. During the telephone interview, respondents would use the calendar as a visual recall aid.

So far, we are not aware of any computerized calendar instruments, which are self-completed by respondents. This is likely to be due to the fact that the current computerized versions, are quite complicated to fill out, and cannot be used without training. Nonetheless, at least one electronic EHC is currently used as a separate recall aid in a self-administered survey (Wiebe and Landis 2000). Before the interview, interviewer and respondent fill out the EHC together. It is subsequently displayed on a laptop screen, during both, the CAPI and a short ACASI section of the interview.

From what is known so far about the relative effectiveness of these different calendar applications, it cannot be concluded that one mode of data collection is more suitable for using calendar applications than others. It is also not obvious if those instruments work better as recall aides or as data collection devices. There are some indications that respondents might be more motivated to make active use of the calendar if they can see the instrument during the interview. This suggestion is based on the observation that respondents prefer calendar techniques to regular question-list surveys in personal interviews (Freedman et al. 1988) but

are largely indifferent to them in CATI surveys (Belli et al. 2001), where they cannot make use of their visual qualities.

4 Effects of calendar methods on data quality

4.1 Evaluations of data quality

As early as in the late 1960s Balán et al. (1969) concluded, in what was probably the first (non-experimental) evaluation of calendar methods, that the calendar instrument had the following advantages over traditional question-list surveys:

1. It improved the completeness of reports by enabling the interviewer to detect ‘gaps’ in the data provided by the respondent.
2. Inconsistencies in the account could be detected by the interviewer or by the respondent himself. The respondent could then correct his original account.
3. It facilitated recall for distinct events, by displaying those events as part of a sequence. This (supposedly) lead to a reduction of omissions.
4. It improved timing of recalled events by allowing the respondent to relate events and dates from different life domains to each other.

Although the study by Balán and his colleagues did not have an experimental design, their observations are still valid today. The expected positive effects of calendar methods on completeness, consistency, recall and timing—as well as the implied effective mechanisms of the calendar—are main issues in evaluations of calendar methods.

Over the years several authors, though not explicitly referring to the Balán-study, have tested one or more of the four statements mentioned above. Many have also included more general observations about the data collection process, such as experiences with different modes of data collection (see previous section), respondent-interviewer rapport, and consequences for the duration of the interview. The body of research on calendar methods also includes a few psychometric studies on reliability and/or validity of data collected with calendar instruments.

Our review of the methodological literature reveals that the quality of the data collected with calendar instruments has been evaluated in multiple ways. The studies can be grouped into three categories:

1. Comparisons of calendar data with similar data collected with more traditional questionnaires (in a split-ballot or otherwise), but without the availability of an external standard of comparison (Becker and Diop-Sidibe 2003; Becker and Sosa 1992; Engel et al. 2001b; Goldman et al. 1989; Yoshihama et al. 2005);
2. Studies in which the agreement between data collected with a calendar instrument and external data sources is measured, but no comparisons are made with regular questionnaires. External data sources include physicians’ records (Rosenberg et al. 1983; Wingo et al. 1988) or reports from earlier waves of longitudinal surveys (Freedman et al. 1988).
3. Experimental studies, which combine the two approaches. Here, the authors assess the agreement between calendar data and external data, and also include a control condition, in which a traditional questionnaire is used (Belli et al. 2001; Van der Vaart 1996, 2004; Van der Vaart and Glasner 2005).

First, we will turn to the first group and present some findings based on indirect comparisons between calendars and traditional questionnaires, after that results from the ‘agreement studies’ (group 2 and 3) will be presented.

4.2 Indirect comparisons between calendar data and regular survey data

The focus of the first group of studies is mainly restricted to indirect measures of data quality, in particular consistency of the data based on logical arguments (e.g., in most societies, there should be no overlap between marriages), completeness of the data (e.g., the detection of “gaps” in employment histories) and patterns in recalled dates (such as the use of prototypical values, i.e., “heaping”). Since these studies do not include an external standard of comparison, they cannot provide direct evidence for the superiority of calendar data in terms of accuracy. However, as will be illustrated below, they do provide some indications that the calendar method overall performs better in collecting recall data than the traditional question-list.

A split-ballot comparison between a calendar method and a traditional questionnaire in a fertility study (Becker and Sosa 1992) indicated that the use of the calendar resulted in more consistent reports. It demonstrated that the calendar method resulted in less superposition of (supposedly) mutually exclusive behaviors: significantly less overlap of advanced pregnancy and contraception use was reported in the calendar condition (1.3%) than in the traditional interview (10.3%). Also supporting positive calendar effects, an interaction was found between the recency of the behavior and the effect of the calendar (Goldman et al. 1989). Goldman and her colleagues note that the calendar instrument was especially effective in enhancing recall of contraceptive use in the beginning of the reference period. A similar effect was found in a study of domestic violence victimization (Yoshihama et al. 2005). The results indicate that higher lifetime victimization rates in the calendar condition were caused by the fact that more respondents reported incidents, which took place in the distant past.

Studies that evaluated retrospective data in terms of completeness mostly concluded that the calendar method performs better than the traditional question-list. Calendars were found to be more helpful in reducing the amount of time unaccounted for in the respondent’s life course (Engel et al. 2001b; Goldman et al. 1989). This reduction is likely to be due to the visual nature of the calendar, which makes it easier for the interviewer to detect those left-out periods and ask the respondent about them (Balán et al. 1969). Overall, calendars appear to result in higher numbers of reported events and episodes, which is usually interpreted as a positive effect (Becker and Sosa 1992; Engel et al. 2001b).

Regarding the heaping of reported event dates—which occurs when respondents report prototypical values (e.g., courses starting in September, or “the accident happened two year ago”) instead of the actual values—only few studies are known to evaluate calendar effects (see also the next section). In an experimental evaluation Goldman et al. (1989) found that the calendar method significantly reduced heaping in reports of contraceptive use. While in the traditional questionnaire condition a disproportionate number of women rounded durations to prototypical values of 6, 12, 24, 36, and 48 months of use, this hardly occurred in the calendar condition. It should be noted however, that this difference was probably enhanced by the coding protocol. While in the questionnaire condition, interviewers could record durations in either months or years; in the calendar condition, interviewers were instructed to always code durations in months.

4.3 Agreement between calendar data and external sources

The second and third group of studies focus on direct assessments of agreement between the recalled information and the external information: in particular concerning the number of events, their characteristics and the duration or dates of events. Some authors turned to data sources such as doctors’ records (Rosenberg et al. 1983), purchase records (Van der Vaart and Glasner 2005) or population registers (Auriat 1993) to validate the retrospective

reports. In the absence of this type of validating information, authors compared calendar data with respondents' earlier (concurrent) reports from the same longitudinal study (Belli et al. 2001; Freedman et al. 1988; Van der Vaart 1996). It can be argued that comparisons of the latter type are an assessment of (test–retest) reliability rather than of validity (Dex 1995). Nevertheless, it seems reasonable to assume that the amount of error is smaller in concurrent than in retrospective reports, since the former are less affected by memory bias. As illustrated below, the results of these both types of studies generally suggest that the calendar method has beneficial effects on data quality.

4.4 Non-experimental validation studies

While non-experimental agreement studies do not compare the performance of the calendar method to the performance of other methods, they do give an indication of the quality of calendar data. In this line, Rosenberg et al. (1983) performed a record check study, which did not include a comparison with another type of questionnaire. Using doctors' records as validation measures the authors report an agreement of 90% between the calendar data and the records for month-specific use of oral contraceptives. The mean duration of the reference period was 33 months. The agreement between physicians' records and self-reports decreased when brand and dose of contraceptive were also considered.

High levels of data quality were also reported in non-experimental longitudinal studies. In their evaluation of calendar questionnaires Hoppin et al. (1998) report very high test–retest reliability of pesticide use when respondents were contacted by telephone one to three weeks after the original interview. A more detailed study of test–retest reliability of the calendar method—the time between the interviews being eight to fourteen months—resulted in very high agreement for reported life event anchors such as marriages, or immigration (Engel et al. 2001a). Freedman et al. (1988) compared respondents' self-reports from two waves (1980 and 1985) of a longitudinal study. In the 1985 wave, a calendar instrument was used. The authors found an 87% agreement between school attendance reported concurrently in 1980 and retrospectively in 1985. Part-time school attendance was remembered less well than either full-time attendance or no attendance. Responses about work in 1980 were less consistent. Here, the agreement between waves was 72%. The general tendency to underreport unemployment in retrospective surveys was not fully compensated for by the calendar.

Thus, several life course studies that applied event history calendars report relatively high correspondence between retrospective calendar data and matching responses or collateral reports obtained beforehand. Similar results are found in small-scale medical studies on health timelines (e.g., Searles et al. 2000). Although these results suggest positive effects of the calendar procedures on recall accuracy, they lack an experimental design: since there is no control condition, it has not been demonstrated whether these results would have been different in a study without aided recall procedures.

4.5 Quasi-experimental studies

Only three studies so far (Belli et al. 2001, 2004; Van der Vaart 1996, 2004; Van der Vaart and Glasner 2005) have combined the approaches depicted above with an experimental design. The authors conducted split-ballot experiments in which they used calendar instruments in one condition and traditional questionnaires in the other condition. Belli et al. (2001) and Van der Vaart (1996) then validated the data from the two conditions with earlier reports from the longitudinal studies. Van der Vaart and Glasner (2005) used purchase records as

validation data. Given the relevance of these studies we will discuss their results in detail below.

In the 1996 study [Van der Vaart \(1996, 2004\)](#) developed and tested a calendar method (in these studies called a 'timeline') that was filled out by the respondents during a face-to-face interview and was subsequently used as a visual recall aid. The calendar was tested in a field experiment on educational careers during the second wave of a longitudinal social survey, comparing the retrospective reports with reports during the first wave four years before (the recall period was 4–8 years). As compared to the regular questionnaire procedure, adding the calendar enhanced data quality with respect to the number of educational courses followed, the starting year of the courses, and the entire sequence of types of courses taken. Although the calendar reduced recall error in the dates of courses, it did so for absolute error only: it did not affect telescoping (i.e., the direction of the net error in dates) and neither did it diminish the heaping effect in reported dates. The calendar was shown to be most effective if respondents had to perform relatively difficult retrieval tasks in terms of recency, saliency, and frequency of the target behaviour (e.g., for respondents who had followed a great number of courses).

Comparable results were found by [Belli et al. \(2001, 2004\)](#) who evaluated an event history calendar by means of a field experiment integrated into a longitudinal household study on social and economic behaviours. All interviews in this study were conducted via telephone in 1998. Respondents were asked for retrospective reports on the number and the duration of events that occurred in 1996. The quality of the 1998 reports using either a calendar—that was visible to the interviewer only—or a question-list, was assessed using data from the same respondents collected one year earlier on events in 1996.¹ Compared to the question-list survey the calendar instrument resulted in significant difference scores, indicating positive effects, for three out of six topics: the number of moves, the number of jobs and the number of persons entering the residence. No differences in data quality were found regarding the number of persons leaving the residence, whether having received children aid and whether having received food stamps. Regarding four out of six continuous measures the calendar method led to significantly higher correlations with the 1996 reports than the question-list. This applied to 'income' (a) and the durations of periods 'being unemployed' (b), periods 'missing work due to illness' (c) and periods 'missing work due to illness of others' (d) No differences in correlations were found for the duration of periods 'working' (e) and periods 'on vacation' (f). In spite of the effects on correlations, hardly any differences in mean errors were found between both conditions.

Finally, the experimental record check study by [Van der Vaart and Glasner \(2005\)](#) generally confirmed the findings of both field experiments presented above. In this study a calendar was employed as a visual aid for respondents in a telephone survey. Unlike most calendar instruments used in the social sciences, this calendar aimed to enhance the recall of singular events (the purchase of pairs of glasses) instead of episodes. The respondents' retrospective reports about a recall period of over 7 years, were compared to database information on three issues: the *price* and the *date* of the latest purchase of pairs of glasses and the *number* of purchases. Hardly any effects could be established regarding the *number* of purchases due to a lack of variation in this measure. Regarding both the *price* and the *date* of the purchase this study demonstrated that:

¹ There were two versions of the question-list: one in which the respondents were asked first about 1996 events and then about 1997 events and a second version that asked first about 1997 and then about 1996. Both question-lists were taken together as the 'standard condition'. Since the 1997 version led to worse data quality than the 1996 version, the design of the study might have weakened the question-list condition and thereby might have caused some overestimation of the positive effects of the calendar method.

- (a) The calendar had positive effects on recall accuracy, although it did not affect telescoping (net error in dates);
- (b) A more difficult recall task—in terms of the saliency and recency—led to greater recall errors;
- (c) Employing the calendar was especially effective in enhancing recall accuracy when the respondents' recall task was relatively difficult, that is: for less salient and less recent purchases.

As will be discussed in more detail below, a downside of this procedure was that the response rate in the calendar condition was quite low. Sending respondents the calendar instrument beforehand probably increased the risk of refusal.

Overall the results of these experimental studies—that compared the calendar method and the question-list method by using external validating data—are mixed but quite promising. They demonstrated that the calendar method exerted positive effects on recall accuracy for different types of data and never led to worse data quality.

5 Operational costs of calendar procedures: fieldwork and sampling

While calendar applications may have beneficial effects on data quality, it is also important to judge their operational costs in terms of interview time, data entry, interviewer training, and so forth, and the effects on sampling. Regarding these issues several potential disadvantages of the use of calendar instruments emerged from the literature.

5.1 Interview time and data entry

One rather consistent finding is that calendar instruments take longer to administer than traditional questionnaires. [Engel et al. \(2001b\)](#) report that their personal interviews, in which the calendar method was used for data collection, took approximately twice as long as the traditional question-list interviews. In a study in which the calendar was used as a recall aid, the face-to-face interviews on average took 12 min longer in the calendar condition, the mean length of all interviews being about 2 h ([Van der Vaart 1996](#)). When a similar calendar application was sent to respondents beforehand and subsequently the interview was held by telephone, the interview took only marginally longer than in the control condition without calendar ([Van der Vaart and Glasner 2005](#)). The same applied to relatively simple calendar instruments for recording fertility-related events ([Becker and Sosa 1992](#); [Goldman et al. 1989](#)). It must be noted, however, that in at least one of the studies, reference periods differed significantly between instruments, probably leading to an under-estimate of duration of the calendar interview.

While a CATI calendar application did not increase total interviewing time, data entry did take significantly longer than for regular questionnaires ([Belli et al. 2001](#)). Freedman and her colleagues (1988) made a similar observation and noted that their calendar data was more difficult and expensive to code than a conventional questionnaire.

5.2 Interviewer training

Although it seems plausible that calendar methods would require more interviewer training than standardized question-lists, this finding did not emerge from the few studies that report about this issue. The extra amount of interviewer training required for administering calendar instruments varies widely by study. Some studies report that interviewers, who had to

use a calendar received the same amount of training as interviewers, who administered the traditional questionnaire (Belli et al. 2001; Van der Vaart 1996; Van der Vaart and Glasner 2005). However, other authors state that training time was tripled compared to earlier waves of their survey, when no calendar instruments were used (Freedman et al. 1988). It seems to depend very much on the individual study, how much extra training is required in order to make sure that the interview is administered properly.

5.3 Non-response

Only a few studies give some indications concerning the impact that calendar methods may have on response rates. Furthermore, the exact impact can be hard to determine since calendars sometimes are administered as just one of the elements of the data collection (for example during one section of an interview). This latter situation is found in the study by Van der Vaart (1996), where the calendar was part of the second wave of a large-scale data collection process; apparently the calendar had no effects on the response rate (71% in this second wave). The study by Belli et al. (2001) involved a separate data collection for the calendar condition and the question-list condition and the response rates—that could be established independently—appeared to be equal (84%) in both conditions. Also in a study in which a calendar was used as a visual recall aid during a personal interview, no effect on the non-response was found (Kominiski 1990). However, the authors note that interviewers and respondents were not required to actively use the calendar.

A different picture emerges from a split-ballot consumer survey in which the calendar—that was used as a recall aid during the interview—was sent to the respondents prior to the telephone interview (Van der Vaart and Glasner 2005). This procedure led to a response rate that was substantially lower in the calendar condition (39%) than in the regular condition (67%). Quite possibly respondents were scared off by the extra work they had to put in when a calendar instrument was used. Yoshihama et al. (2005) found an even larger difference in response rates between the calendar condition and the question-list conditions. In spite of the fact that both samples were selected from one (listed) operational population, using the same criteria, the response rate was only 18% in the calendar condition versus 78% in the control condition. All in all it seems clear that more studies are needed to determine potential effects of the use of calendar methods on response rates.

6 Evaluations of the interviewing process

Since the calendar method affects the questioning procedure and the tasks of both the interviewer and the respondent substantially, it probably has motivational effects next to memory effects. However, whether these motivational effects prove to be positive or negative is hard to say. On the one hand, a calendar procedure may have positive motivational effects since it is a less well-known method and asks for a more active approach by the interviewer and/or respondent. This may suggest to the interviewer and respondent that their task is important and needs precision, as is the case for aided recall methods in general (Sudman and Bradburn 1983). But on the other hand, these very same characteristics might also cause fatigue, meaning that the task may be too burdensome for the interviewer and/or the respondent (Billiet et al. 1984). While researchers usually appreciate the positive influence of calendar instruments on data quality, they also note that the application of a calendar method often increases the length of the interview and sometimes of the interviewer training time (see the previous section). However, reports of experiences with calendars as instruments for data

collection are generally quite positive. This is especially true for interviewers' and respondents' subjective evaluations of the instrument (Martyn and Martin 2003).

Results of evaluation surveys among interviewers, who worked with calendar instruments, suggest that these instruments are perceived as being interesting to work with (Belli et al. 2004; Freedman et al. 1988). Among others, the interviewers' preference for calendar instruments over regular questionnaires is attributed to the fact that inconsistencies in the data can be removed while the respondent is still available for clarification (Balán et al. 1969; Belli et al. 2001; Goldman et al. 1989). Generally these instruments are also perceived to yield better data quality than do standard questionnaires. Interviewers feel that calendar instruments help respondents to understand questions better (Belli et al. 2001). Interviewers also think that calendars make the interview "more enjoyable" for the respondent (Freedman et al. 1988). Similarly interviewers found that the calendar helped them discuss sensitive issues with their respondents (Martyn and Martin 2003).

Interviewer perceptions that respondents prefer calendar instruments to traditional questionnaires are confirmed by respondents' feedback on the interviewing process. In the study by Martyn and Martin (2003), mentioned above, respondents reported that the calendar made it easier for them to discuss sensitive issues with the interviewer because they could refer to the information they had written down in the calendar. In addition they stated that the calendar did 'jog their memory' by helping them relate events from several life domains to each other. Respondents are also reported to enjoy filling in the calendar (Hoppin et al. 1998). In a comparison of a calendar-based questionnaire with a traditional questionnaire, respondents appeared to be more patient and cooperative in the calendar condition (Engel et al. 2001b). They were more concerned with data quality when a calendar was used, and sometimes even asked for a copy of the completed calendar to take home with them. Caspi and his colleagues (1996) note that, in a pilot study, respondents actively corrected the information in the interviewer-administered paper-and-pencil calendar. Literature on the use of satisfying strategies by respondents suggests that this higher involvement with the interviewing process and objective should have a positive effect on data quality (e.g., Holbrook et al. 2003).

On the whole, the respondent feedback about calendar instruments is often positive. This is especially true for interviews in which respondents can see the calendar and use it as a visual recall aid (e.g., Engel et al. 2001b; Hoppin et al. 1998). When calendars are used mainly as a tool for the interviewer, interviewer evaluations of the procedure tend to be more positive than respondent evaluations (Belli 2000).

7 Conclusion and discussion

This review illustrates that applications of calendar instruments in social research have been growing rapidly during the last decade. Calendar instruments are used in a wide variety of research fields and with very diverse populations. The instruments are used in personal as well as in telephone interviews, and they serve either purely as a recall aid or as an instrument of data collection. In some studies respondents filled in calendars according to written instructions and without the assistance of an interviewer. Computerized versions of calendar instruments are available for both, CATI and CAPI applications, but not (yet) for self-completed questionnaires.

Methodological studies on the use of calendar techniques show mixed results with regard to effects on data quality. Effects are found for some issues, but not for others. Sometimes those effects are strong, mostly they are modest. Most calendar instruments are found to increase the completeness of respondents' accounts. This is especially true for the reduction

of ‘gaps’, i.e., time unaccounted for. Often also beneficial effects are reported on the accuracy of the number and characteristics of reported events. From the empirical evidence we have, it can be concluded that calendars do also enhance consistency of responses. Additionally, they are reported to lead to a reduction of dating error, though results on telescoping and heaping are mixed. Furthermore, several studies report interaction effects of the calendar application with the difficulty of the recall task, indicating that calendars might be especially helpful for recalling less recent, more frequent and less salient events. However, up to the present the number of methodological studies is limited and most of the supposed beneficial effects of the calendar reported in this review need more empirical evidence.

The potential operational costs, in terms of increased interview duration and additional interviewer training, as well as consequences for sampling, are not clear yet: both positive and negative consequences were reported. The interviewer response to calendar instruments is generally good, while respondent evaluations seem to depend on the degree to which respondents can see and experience the calendar. This ‘user evaluation’ is important. Since working with calendar instruments requires some extra effort, it is crucial to keep up interviewer and respondent motivation. Overall, it is encouraging that the operational problems that arose in this review do not appear to be insurmountable. Again, more research is needed and then may lead to ‘best practices’.

We should be aware of the fact that applying calendar procedures can also be counterproductive. They may create consistency and completeness in the data that does not represent higher validity but biased reconstructions instead. This artifact may arise especially if respondents’ behavior is less consistent than is assumed by the researcher. It is possible, for example, that women use contraceptive methods during pregnancy (Becker and Sosa 1992), or that somebody holds a fulltime and a part-time job at the same time. Likewise, in one culture certain events or states might be mutually exclusive and thus reflect response inconsistency—e.g., multiple simultaneous marriages—while they are not in another culture (Axinn et al. 1999). While in many cases removal of such inconsistencies will result in more valid data, it will lead to error in other cases. Correspondingly, when it comes to ‘time unaccounted for’ in the respondent’s life history, asking the respondent to ‘fill in the gaps’ might not always be a good decision. When the gap stands for an episode that the respondent cannot recall during the interview, he or she might stretch previous or subsequent episodes in order to reduce the amount of time unaccounted for, leading to decreased data quality (Auriat 1993).

The literature review revealed that the growing interest in calendar methods resulted in many different applications with many different names, which suggests that researcher do not take enough advantage of each other’s efforts. It also appeared that most applications are characterized by a modest theoretical foundation, irrespective of the apparent fact that calendar methodology may utilize organizing principles of autobiographical memory (Belli 1998; Belli et al. 2001). Theories and insights from cognitive science and related fields (see Tourangeau et al. 2000) can be further employed to formulate a theoretical framework on recall bias and calendar techniques in social survey. It is clear that a calendar procedure has to be attuned to the subject matter of the specific survey. Different topics (e.g., educational history versus alcohol consumption) and different populations also entail differences in the requested information and thus different recall tasks. Computer-assisted applications may be helpful to adjust the procedures to a certain (group) profile that is derived from the respondent’s answer’s to earlier questions from the questionnaire, like has been done in ‘pre-loading’ electronic questionnaires (Hoogendoorn 2004).

However, from the review it appears that it often remains unknown and unspecified, which design characteristics of calendars have which kind of effects on recall accuracy. Therefore a main line of future methodological research would be to perform experimental studies that

can shed light on this relationship. Those studies might address topics such as the optimal length of the reference period, number of domains included for maximum effects, effectiveness of different kinds of landmarks, or the effectiveness of the calendar together with different modes of data collection. The issue of non-response also remains very important. Apart from that, further research is needed regarding the roles of the interviewer and respondent. This includes questions like: when should the calendar be used as a questionnaire in itself and when as a separate recall aid? How active and/or steering should the role of the interviewer be? In what way should the interviewer probe and/or help the respondent fill out the calendar?

In sum, our proposal for future research involve a research program that includes at least the following variables:

1. Independent variables: properties of calendar techniques, such as types of cues and landmarks, the bounding of reference periods, and the graphical design.
2. Conditional variables: the difficulty of the recall task (like the length of the recall interval), group differences (e.g., age groups) and modes of data collection.
3. Dependent variables: data quality of responses to factual, retrospective questions on the occurrence, properties, dates and duration of life events in different life domains.

Both small-scale experiments and experimental studies in large-scale surveys are required in order to provide information on the working and effectiveness of calendar applications. In order to further develop and adjust the rationale of the calendar methodology, studies are needed that explicitly test the theoretical ideas that underlie a specific calendar application. It is obvious that many fields might benefit from such methodological studies into calendar procedures. Calendar instruments emerge from the literature as promising methods to enhance retrospective reports. A more systematic approach in the development of calendar methodology might increase their relevance and effectiveness substantially.

Acknowledgements This study is part of the research project “Reconstructing event histories in standardized survey research: cognitive mechanisms and aided recall techniques” which is supported by the Netherlands Organization of Scientific Research, Grant Number 400-03-331.

Appendix

A truncated example of a filled out calendar (In: [Van der Vaart and Glasner 2005](#)).

| 2003 | January | February | March |
|-----------------------|-------------------------|---------------------|--------------------|
| <i>interview</i> | | | |
| 1. Age | | 48 | |
| 2. City and street | A'dam, Parkstr — — — — | — — — — — — — — | A'dam, Churchstr → |
| 3. Domestic situation | — — partner and son — — | — — partner — — — — | — — — — — — — — → |
| 4. Jobs and courses | — — ING bank — — — — | — — — — — — — — | — — — — — — — — → |
| 5. Personal landmarks | | Indonesia | robbery |

References

Auriat, N.: My wife knows best: A comparison of event dating accuracy between the wife, the husband, the couple, and the Belgium Population Register. *Public Opin. Q.* **57**(2), 165–190 (1993)

Axinn, W.G., Pearce, L.D., Ghimire, D.: Innovations in life history calendar applications. *Soc. Sci. Res.* **28**, 243–264 (1999)

- Balán, J., Browning, H.L., Jelin, E., Litzler, L.: A computerized approach to the processing and analysis of life histories obtained in sample surveys. *Behav. Sci.* **14**(2), 105–120 (1969)
- Becker, S., Diop-Sidibe, N.: Does use of the calendar in surveys reduce heaping?. *Stud. Fam. Plan.* **34**(2), 127–132 (2003)
- Becker, S., Sosa, D.: An experiment using a month-by-month calendar in a family planning survey in Costa Rica. *Stud. Fam. Plann.* **23**(6), 386–391 (1992)
- Belli, R.F.: 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 (1998)
- Belli, R.F.: Computerized Event History Calendar methods: Facilitating autobiographical recall. American Statistical Association, Alexandria (2000)
- Belli, R.F.: Improving the quality of retrospective reports: Calendar interviewing methodologies. Paper presented at the Sixth international conference on logic and methodology, Amsterdam (2004)
- Belli, R.F., Lee, E.H., Stafford, F., Chou, C.-H.: Calendar and question-list survey methods: association between interviewer behaviors and data quality. *J. Off. Stat.* **20**(1), 185–218 (2004)
- Belli, R.F., Shay, W., Stafford, F.: Event history calendars and question list surveys: a direct comparison of interviewing methods. *Public Opin. Q.* **65**, 45–74 (2001)
- Billiet, J., Loosveldt, G., Waterplas, L.: Het survey-interview onderzocht: Effecten van het ontwerp en gebruik van vragenlijsten op de kwaliteit van de antwoorden. Sociologisch Onderzoeksinstituut, Leuven (1984)
- Brückner, E., Mayer, K.U.: Collecting life history data: Experiences from the German life history study. In: Giele, J.Z., Elder, G.H. (eds.) *Methods of Life Course Research: Qualitative and Quantitative Approaches*, pp. 152–181. Sage, Thousand Oaks (1998)
- Caspi, A., Moffitt, T.E., Thornton, A., Freedman, D., Amell, J.W., Harrington, H. et al.: The life history calendar: A research and clinical assessment method for collecting retrospective event-history data. *Int. J. Meth. Psychiatr. Res.* **6**, 101–114 (1996)
- Colt, J.S., Engel, L., Keifer, M., Thompson, M.L., Zahm, S.: Comparability of data obtained from migrant farmworkers and their spouses on occupational history. *Am. J. Indus. Med.* **40**, 523–520 (2001)
- Conway, M.A.: A structural model of autobiographical memory. In Conway, M.A., Rubin, D., Spinnler, H., Wagenaar, W.A. (eds.) *Theoretical Perspectives on Autobiographical Memory*. Kluwer Academic Publishers, Dordrecht (1992)
- Dex, S.: The reliability of recall data: a literature review. *Bulletin de Methodologie Sociologique* **49**, 58–89 (1995)
- Engel, L., Keifer, M., Thompson, M.L., Zahm, S.: Test–retest reliability of an icon/calendar-based questionnaire used to assess occupational history. *Am. J. Indus. Med.* **40**, 512–522 (2001a)
- Engel, L., Keifer, M., Zahm, S.: Comparison of a traditional questionnaire with an icon/calendar-based questionnaire to assess occupational history. *Am. J. Indus. Med.* **40**, 502–511 (2001b)
- Freedman, D., Thornton, A., Camburn, D., Alwin, D., Young-DeMarcco, L.: The life history calendar: A technique for collecting retrospective data. *Sociol. Methodol.* **18**, 37–68 (1988)
- Giele, J.Z., Elder, G.H.: *Methods of Life Course Research: Qualitative and Quantitative Approaches*. Sage, Thousand Oaks (1998)
- Goldman, N., Moreno, L., Westoff, C.F.: Collection of survey data on contraception: An evaluation of an experiment in Peru. *Stud. Fam. Plann.* **20**(3), 147–157 (1989)
- Holbrook, A.L., Green, M.C., Krosnick, J.A.: Telephone versus Face-to-face interviewing of national probability samples with long questionnaires: comparisons of respondent satisficing and social desirability response bias. *Public Opin. Q.* **67**(1), 79–125 (2003)
- Hoogendoorn, A.W.: A questionnaire design for dependent interviewing that addresses the problem of cognitive satisficing. *J. Off. Stat.* **20**(2), 219–232 (2004)
- Hoppin, J., Tolbert, P., Flagg, E., Blair, A., Zahm, S.: Use of a life events calendar approach to elicit occupational history from farmers. *Am. J. Indust. Med.* **34**, 470–476 (1998)
- Kominski, R.: The SIPP event history calendar: Aiding respondents in the dating of longitudinal events. In: Association, A.S. (ed.) *Proceedings of the Section on Survey Research Methods*. American Statistical Association, Alexandria (1990)
- Martyn, K.K., Martin, R.: Adolescent sexual risk assessment. *J. Midwifery Womens Health*, **48**(3), 213–219 (2003)
- Pebley, A.R., Sastry, N.: *The Los Angeles Family and Neighborhood Survey: Field Interviewer Manual*. RAND (2004)
- Reimer, M.: *Collecting Event History Data About Work Careers Retrospectively: Mistakes that Occur and Ways to Prevent Them*. Max Planck Institute for Human Development, Berlin (2004)
- Rosenberg, M., Layde, P., Ory, H., Strauss, L., Rooks, J., Rubin, G.: Agreement between women's histories of oral contraceptive use and physician records. *Int. J. Epidemiol.* **12**(1), 84–87 (1983)

- Schwarz, N., Sudman, S.: *Autobiographical Memory and the Validity of Retrospective Reports*. Springer-Verlag, New York (1994)
- Scott, J., Alwin, D.: Retrospective versus prospective measurement of life histories in longitudinal research. In: Giele, J.Z., Elder, G.H. (eds.) *Methods of Life Course Research: Qualitative and Quantitative Approaches*. Sage, London (1998)
- Searles, J.S., Helzer, J.E., Walter, D.E.: Comparison of drinking patterns measured by daily reports and timeline follow back. *Psychol. Addict. Behav.* **14**(3), 277–286 (2000)
- Smith, J.P., Thomas, D.: Remembrances of things past: Test–retest reliability of retrospective migration histories. *J. R. Stat. Soc. Ser. A* **166**(1), 23–49 (2003)
- Sobell, L.C., Sobell, M.B.: Alcohol consumption measures. In: Allen, J.P., Wilson, V.B. (eds.) *Assessing Alcohol Problems: A Guide for Clinicians and Researchers*, 2nd edn., pp. 75–99. National Institute on Alcohol Abuse and Alcoholism, Bethesda, MD (2003)
- Sobell, L.C., Sobell, M.B., Leo, G.I., Cancilla, A.: Reliability of a timeline method: Assessing normal drinkers' reports of recent drinking and a comparative evaluation across several populations. *Brit. J. Addict.* **83**, 393–402 (1988)
- Sudman, S., Bradburn, N.: *Asking Questions: A practical Guide to Questionnaire Design*. Jossey-Bass, London (1983)
- Tourangeau, R., Rips, L., Rasinski, K.A.: *The Psychology of Survey Response*. Cambridge University Press, Cambridge (2000)
- Turner, R.J., Wheaton, B., Lloyd, D.A.: The epidemiology of social stress. *Am. Sociol. Rev.* **60**(1), 104–125 (1992)
- Van der Vaart, W.: *Inquiring into the Past. Data Quality of Responses to Retrospective Questions*. PhD-thesis, VU University Amsterdam, Amsterdam (1996)
- Van der Vaart, W.: The time-line as a device to enhance recall in standardized research interviews: A split ballot study. *J. Off. Stat.* **20**(2), 301–317 (2004)
- Van der Vaart, W., Glasner, T.J.: Enhancing recall accuracy in surveys by cognitively tailored timeline methods: A record check study. Paper presented at the First EASR Conference, Barcelona (2005)
- Van der Vaart, W., Van der Zouwen, J., Dijkstra, W.: Retrospective questions: data quality, task difficulty, and the use of a checklist. *Quality Quantity* **29**(3), 299–315 (1995)
- Weinstock, J., Whelan, J.P., Meyers, A.W.: Behavioral assessment of gambling: An application of the timeline followback method. *Psychol. Assess.* **16**(1), 72–80 (2004)
- Wiebe, E., Landis, K.: 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 (2000)
- Wingo, P., Ory, H., Layde, P., Lee, N.C.: The evaluation of the data collection process for a multicenter, population-based, case-control design. *Am. J. Epidemiol.* **128**(1), 206–217 (1988)
- Yoshihama, M., Gillespie, B., Hammock, A., Belli, R.F., Tolman, R.: Does the life-history calendar method facilitate the recall of domestic violence victimization?: Comparison of two methods of data collection. *Soc. Work Res.* **29**(3), 151–163 (2005)