

## The Integration of a Computer Assisted Interviewing Event History Calendar in the Panel Study of Income Dynamics

Robert F. Belli

University of Nebraska--Lincoln

### Introduction

Prospective panel surveys are valued for their ability to collect life course data for purposes of modeling changes in the human condition. Celebrated as the longest continuing panel survey in the world, the University of Michigan's Panel Study of Income Dynamics (PSID) has been interviewing persons from the same households since 1968. Because information is being collected from the same households across time, a key challenge for the PSID is to maintain consistency in the quality of data and data structure in lieu of changes associated with interviewing costs and best practices in survey interviewing, data collection, and management. To manage costs, the PSID moved from face-to-face interviewing to telephone interviewing in 1973, and interrupted its series of annual waves of interviewing to biennial waves between 1996 and 1998. To keep pace with "best practices," the PSID began to adopt computer-assisted telephone interviewing in 1993, with full implementation in 1994 (previously, paper and pencil instruments were used), and in 2003 integrated the use of a 2-year computerized event history calendar interviewing methodology with the traditional standardized question-list approach. The purpose of this paper is to outline the rationale behind the adoption of the computerized event history calendar, and to highlight the advantages to data quality and completeness through this approach.

### Background

Every prospective study has a retrospective component. As for the PSID, annual interviewing permitted the collection of annual information for the past year, which, depending on the date of interviewing, usually involved retrospective reporting of events that happened at most 1½ years previously. With the transition to biennial interviewing following 1996, study staff and the Board of Overseers became concerned with the

threats to data quality that would be necessitated by increasing the length of the reference period, given what is known by the loss of accuracy in memory reports with increases in retention interval as observed both in experimental psychological research (Linton 1982; Rubin and Wenzel 1996; Thompson, Skowronski, Larsen, and Betz 1996; Wagenaar 1986) and in retrospective reports collected in traditional standardized question-list (Q-list) methods (Cannell et al., 1965; Dugoni, Lee, and Tourangeau, 1996; Duncan & Hill, 1985; Mathiowetz & Duncan, 1988).

*EHCs and Response Accuracy.* The use of EHCs or charts to collect retrospective information has had a wide impact on the fields of population studies and sociology (Axinn et al., 1997; Caspi et al., 1996; Freedman et al., 1988; Furstenberg et al., 1987; Kessler & Wethington, 1991; Lyketsos et al., 1994; van der Vaart, 2004) and their use has demonstrated high quality retrospective reporting. For example, Freedman et al. (1988) compared the retrospective reports of school attendance and weekly work hours in the life history calendar against the concurrent reports gathered with Q-list interviewing that had taken place 5 years earlier. In examining the same month of the prior concurrent reports, 87% of respondents gave identical answers in both of the interviews regarding full-time, part-time, or no school attendance. In another application of a life history calendar, Caspi et al. (1996) found at least 90% agreement between retrospective reports of activities for a given month on the life history calendar and concurrent reports obtained 3-years earlier with living arrangements, cohabitation, schooling, employment, and job training.

Theoretically, EHCs acquire their ability to optimize the quality of retrospective reports because they tap into available idiosyncratic structures in autobiographical memory (Belli, 1998). In contrast to Q-lists, which emphasize structuring the interview through standardized scripted questions, EHCs are visually structured instruments that emphasize the interrelationships of spells of activity in one's autobiographical past by presenting timelines to record what events happened, and when. In order to tap into idiosyncratic autobiographical memory structures, EHCs encourage a flexible style of interviewing in which events from respondents' pasts are used as cues to facilitate the recall of other related events. These cueing mechanisms optimize the ability of respondents to reconstruct an accurate and complete memory of the occurrence and

timing of past events, even after retention intervals of many years. Especially for EHC interviews, two cueing mechanisms, sequential and parallel retrieval, are considerably more dominant in comparison to Q-list interviews. In sequential retrieval, remembering an event in a specific domain, such as a particular employment, assists in the remembering of earlier and later jobs. With parallel retrieval, remembering an event in one domain assists in the remembering of contemporaneous events in other domains, such as remembering a change in addresses assists in the remembering of a change in jobs. In addition to these cueing mechanisms, the theoretical benefits of EHC interviewing assume more effective communication patterns, in comparison to Q-list techniques, as EHCs promote a narrative style of remembering which is compatible with the manner in which autobiographical knowledge is structured (Brown & Schopflocher, 1999; Schank & Abelson, 1995).

*The 1998 PSID Calendar Methods Study.* To address concerns regarding the quality of retrospective reports with a biennial interviewing schedule, the Board of Overseers supported a methodological study, supported by additional funding from the National Science Foundation (SBR-9730297) to assess the quality of retrospective reports concerning events that occurred approximately 1 to 2 years previously. In this methodological study, known as the 1998 PSID Calendar Methods Study, the quality of retrospective reports collected with an Event History Calendar (EHC) methodology were compared with those collected with a Q-list instrument. The PSID Calendar Methods Study was the first direct experimental comparison between event history calendar (EHC) and Q-list methods (see Belli, Shay, & Stafford, 2001).

Method. Interviews were collected in 1998 via telephone on paper with a random subset of respondents and questions from the regular 1997 PSID core survey. Two conditions were tested, a Q-list condition ( $n = 307$ ; 84.1% cooperation rate), and an EHC condition ( $n = 309$ ; 84.4% cooperation rate). Respondents and 20 interviewers were randomly assigned to conditions. All interviewers received general interviewer training plus 15 hours (5 hours for each of three days) of training in their respective methods. The Q-list condition used a traditional standardized survey instrument based on the core PSID questionnaire, a 25-page paper questionnaire with scripted question sequences. Interviewers were asked to optimize retrospective reporting within the constraints of

standardized interviewing techniques. The EHC condition employed paper abstracts from the Q-list method as reference material, but the survey instrument, per se, was one 18"x 28" page; flexibility in using sequential and parallel cues were facilitated by the layout of timelines within domains in the instrument. With the EHC, interviewers were given scripted questions to introduce each domain and were instructed on how to use top-down, sequential, and parallel cues. Interviews were conducted in 1998 about events that occurred during 1996 and 1997.

Results. Using data from respondents collected one year earlier on events reported during the 1997 core PSID as a standard of comparison, the quality of retrospective reports on 1996 events from the 1998 administration of EHC and Q-list interviews was assessed. The results of the experiment supported the notion that EHC methods promote a better utilization of beneficial autobiographical memory processes in comparison to state-of-the-art standardized Q-list methods. Better quality retrospective reporting for residence changes, number of jobs, earned income, weeks unemployed, weeks away from work due to personal illness and illness of another, was found with the EHC interviews for retrospective reports that targeted a reference period that was one to two years ago. There were no substantive differences in interviewing time between methods. Follow-up questionnaires revealed that interviewers enjoyed the EHC interviews more than the Q-list ones, they found the EHC easier to administer, and that they believed that the respondents found the questions on the EHC to be easier to answer. Despite the EHC collecting much more detailed information than the Q-list, respondents reported the same levels of burden for both methods.

The benefits of EHC interviewing in comparison to Q-list methods is illustrated by examining the marginal results of reports for residential moves and number of jobs treated categorically. Consider the proportion of respondents who reported moving at least once during 1996 or early 1997, and the proportions who reported having 0, 1, or 2 or more jobs during 1996, when responding in the 1997 core PSID standard of comparison interviews and in the 1998 experimental interviews. Figure 1 depicts the results for reported moves, using two different standard of comparison assessments. With the yes/no standard of comparison, responses of "yes" or "no" to a question on the 1997 core PSID interview on whether respondents reported having moved since the spring of

1996 was used. With the concurrent reports of address as a standard of comparison, the addresses that were concurrently reported by respondents during the 1996 and 1997 regular annual interviews were used; if the addresses were the same, the respondent was determined to not have moved during the overlap period, and if they were not the same, the respondent was determined to have moved. Figure 2 depicts the proportions of respondents on reports of number of jobs. Both figures illustrate the overall tendency in the experimental conditions to underreport relative to the standard of comparison reports. The figures also show that this tendency to underreport is less pronounced in the EHC condition relative to the Q-list condition.

With reported moves (see Figure 1), and the yes/no responses as standard of comparison, the Q-list condition underreported moves by a level of 6.2% (from 18.6% in the standard of comparison to 12.4% in the Q-list) whereas the EHC condition only led to a 1.6% level of underreporting (from 20.5% to 18.9%). Using the concurrent address reports as a standard of comparison, the data pattern with moves is replicated, with the Q-list condition leading to a 8.1% reduction in reported moves (from 20.7% to 12.6%), and the EHC condition resulting in a 4.3% reduction (from 23.0% to 18.7%).

In reporting the number of jobs held during 1996, the Q-list condition underreported the proportion of respondents who reported having held two or more jobs in 1996 by 9.4% (from 24.4% to 15.0%; see Figure 2), and overreported the proportion of respondents who held one job in 1996 by 8.2% (from 51.8% to 60.0%). Overall, then, there is a disproportionate reduction in the number of jobs reported in the Q-list condition relative to the standard of comparison. These results contrast sharply with the EHC condition, which reproduced the proportions found in the standard of comparison for reports of 0, 1, and 2 or more jobs, almost exactly. This finding is especially crucial given the focus of PSID on labor economics.

Figure 3 illustrates the substantive degree of differences in the correlations between experimental and standard of comparison reports that were found for earned income, number of months employed, number of weeks unemployed, months unemployed, weeks away from work due to personal illness, weeks away from work due to illness of another, and months in receipt of ADC. All of these correlations, with the exception of months of ADC, are significantly stronger in the EHC condition. In total,

there were 23 comparisons between EHC and Q-list measures. Of these, the EHC led to higher quality retrospective reports in 12 comparisons, and the Q-list in only one of them (months of ADC). In comparison to the Q-list, the EHC was found to provide better correspondence with the standard of comparison for whether one moved, the number of jobs, income, the number of months employed, the number of weeks and months unemployed, the number of months out-of-the labor force, the number of weeks missing work due to personal illness and the illness of another, which months employed, and which months unemployed. The advantages that the EHC condition reveal in comparison to Q-list interviewing are even more remarkable when considering that the format of interviewing in the EHC condition was less similar to the standard of comparison than that engendered in the Q-list conditions, as the standard of comparison itself was derived from a Q-list interviewing methodology.

Recently, Belli et al. (2004) have provided insights regarding the advantages of EHC interviews in comparison to Q-list ones by conducting a verbal behavior coding analyses of audiotaped interviews from the 1998 PSID Calendar Methods Study. After identifying verbal interaction patterns, it was found that EHC interviewing was marked by the more frequent occurrence of retrieval cues in interviewer probes and in spontaneous respondent retrieval strategies in comparison to Q-list methods. In addition, EHC interviews demonstrated a higher level of conversational engagement than Q-list methods, as indicated by higher levels of clarification and feedback behaviors. Importantly, the greater frequency of cueing mechanisms was associated with higher levels of data quality, but only in the EHC; the greater frequency of conversational engagement was associated with lower levels of data quality, but only in the Q-list. Differences in data quality between conditions were not attributable to differential condition interviewer effects.

#### The EHC in 2003 PSID Core Interviewing

For the 2003 PSID, a computer-assisted interviewing EHC instrument, copyright to the Regents of the University of Michigan, was developed and integrated with a computerized Blaise® instrument. The instrument development was supported, in part, by the National Science Foundation (SES-0001994). The EHC instrument was largely limited to asking timeline-based questions such as employer names and when household

participants had been working for these employers; the Blaise instrument followed with most of the job details, including for each employer the type of work being performed, how much was being earned, and average hours per week worked. The EHC and subsequent Blaise questions were designed to collect detailed information on the interview year (2003), and the calendar year which preceded the interview year (2002); in addition, for specific variables, the EHC and subsequent Blaise questions collected information of the calendar year that had occurred during the “off-year” (2001), that is, two calendar years previously. The design of this “2-year” EHC had been informed by the programming of a prototype 2-year EHC, by the programming of a two-year EHC for the “Los Angeles Family and Neighborhood Survey,” and by the programming of a life course EHC for the “Computerized Calendar Methods: Health and Economic Measures” project (see, for example, Belli, 2000). Before implementing the EHC into production, a series of preproduction phases were conducted, including pretests and pilot interviewing which tested the entire EHC/Blaise combination.

The EHC was designed to collect information on 5 domains, Landmark Events, Residence, Employment, Not Working, and Time Away. With these domains, respondents were often asked to report the intervals of activity in which start and stop dates were acquired, given that the “third-of-the-month” served as the smallest time unit for timeline data entry. The “third-of-the-month” represents a division of each month into thirds, the beginning, end, and middle of the month. For purposes of interviewing training, intervals of activity were referred to as “spells.” For example, if someone reported working for “Joe and Sons” from the beginning of March to the end of December, 2002, this interval of employment activity would represent a “job spell.”

To facilitate recall of events in the reference period, the Landmark Events sought easily remembered major events that occurred during 2002 and 2001, to serve as contemporaneous anchors to assist in the remembering of residential and labor transitional events that were queried later in the interview. The Residence domain sought the addresses of places that the head of household (head) had been living from January 2001 until the date of interview. The Employment domain asked about all of the different employers of the household participants (both head and wife/”wife”) from January 2001 until the date of interview. As for the Not Working domain, spells of

unemployment, and being out of the labor force, were queried concerning both head and wife/"wife" for the calendar years of 2002 and 2001. Finally, with the Time Away domain, the statuses of both the head and wife/"wife" were queried with regard to missing work due to another person's illness, one's own illness, being on vacation, on strike, or temporarily laid off.

As the PSID is a household interview, the exact domains that were covered during an interview was dependent on who served as a respondent for the household. If the head of household or wife/"wife" served as the respondent, the landmark events domain was asked, otherwise, the landmark events domain was skipped. Because the landmarks events domain was designed to provide parallel anchors to the collection of residential and labor histories of the head and wife/"wife," any landmark events provided by a person other than head and wife/"wife" is potentially unrelated to their respective residential and labor events. If the head served as the respondent, only the head's residential history was collected; similarly, if the wife/"wife" served as the respondent, only the wife/"wife's" residential history was collected. As the designation of an individual as a wife/"wife" would have been made by the 2001 interview, the most likely scenario is that the residential histories of head and wife/"wife" would be identical, or nearly so.

Figure 4 is a screen capture of the appearance of the EHC when collecting information from the Employment domain. As the figure shows, time is represented by calendar year, season, and month. In the data entry window, each month is split into thirds. Among the strategies that are available to assist respondents to remember the timing of transition in events, such as starting a new job, interviewers are trained to narrow the timing by first asking for calendar year, then season, then month, then third-of-a-month. The computer screen is divided into three sections. The *top section* provides summary information regarding information from key domains that have already been collected during the interview for purposes of facilitating parallel retrieval. To highlight available specific retrieval cues, the interviewer places the cursor on a specific calendar node; in this case note the reference to the respondent having lived at "321 Real Nice Place." The *middle section* permits time line entries of the domain that is currently queried, such as entering an interval of having worked for "Joe and Sons" and "Bill's



Landscaping “. The *bottom section* facilitates the entry of specific data that corresponds to each spell, such as the name of an employer. Not illustrated here, the program also permitted interviewers to view introductory scripts to each domain.

Within the middle and top sections, the color red is used to highlight potential inconsistencies (see Figure 4). Note that within the “Not working” timeline in the top and middle sections, and with the timeline entry in the middle section for “Joe and Sons,” there is a potential inconsistency in an interval of time in which a respondent reports working and not working at the same time. A special benefit in collecting labor history information with the computerized EHC is that there are programmed checks that are based on resolving inconsistencies and ensuring that timelines are completed to their entirety. Inconsistencies include reports of working and not working at the same time, and of having time away from work when there was no interval of employment. As for the completeness of the data that is collected, an exhaustiveness criterion stipulates that there cannot be any gaps in the labor history; an individual is expected to be classified either as working, or unemployed, or out-of-the-labor-force at any point in time.

Every attempt has been made to design the EHC as to maintain consistency and continuity in the structure of the panel data. At times, in order to maintain data consistency while ensuring a smooth flow to the interviewing, Q-list questions were directly integrated into the EHC itself. Figure 5 illustrates the use of the traditional employment status questions that have been used in prior panel waves before the employment timelines are displayed. The collection of intervals with start and stop dates to the thirds-of-a-month facilitates the computation, through algorithms, of weeks employed, unemployed, out-of-the-labor-force, and missing work due to another’s or one’s own illness, and due to being on vacation, on strike, or temporarily laid off.

For purposes of comparison, the corresponding questions and variables covered by the 2003 EHC implementation as they had appeared in 2001 are summarized and contrasted in Table 1. Italicized entries indicate 2001 variables that are not available in 2003, and center on the decision for 2003 not to collect information on position changes within the same job. Bold entries indicate 2003 information that had not been previously collected, and center on collecting data on all residence changes (not just the most recent move), greater precision in the dating of labor history statuses (to thirds-of-a-month

instead of to months), gaining more complete information on weeks worked and hours per week for each job worked (not just “main” jobs), and a greater amount of labor history information for the calendar year that occurred 2-years prior to the interview year (T-2 information).

#### The Potential Future of the EHC in the PSID

Because of the ability of EHCs to improve the quality of retrospective reports, the EHC methodology is planned for continuing implementation as a bridge between waves in the foreseeable future. In addition, recently collected evidence from the Health and Economic Measures (HEM) Lifecourse Validation Study project (jointly funded by NIA and NICHD – 1R01AG17977) demonstrates that EHC interviewing leads to better quality *life course* retrospective reports in comparison to Q-list methods. At present, the PSID implements a Q-list methodology to collect life course retrospective reports during intake interviews of new households and in collecting information from new heads or wives/”wives.” In order to improve the quality of these reports, the implementation of a life course EHC may be pursued in future waves of the PSID.

Interviews from the HEM Lifecourse Validation Study were collected during three months of interviewing between July and September of 2002 via telephone, with a random subset of respondents from the regular 2001 Panel Study of Income Dynamics (PSID) core survey. Eligible respondents for the study were PSID study participants since at least 1980, who were respondents for at least ½ of the years they were in the panel, and were at least 45 years of age. There were two conditions, a Q-list condition (N=315, 96% cooperation rate) and an EHC condition (N=311, 93% cooperation rate). Respondents were randomly assigned to one of the two conditions. Twenty-eight interviewers were first matched on experience and then randomly assigned to one of the two conditions. All twenty-eight interviewers received a total of 21 hours of training over three days. The CATI EHC and Q-list were parallel instruments that collected retrospective information for respondents’ entire life courses. Domains that collected retrospective information on residence, marriage and cohabitation, labor, and health histories could be validated from PSID prospectively collected panel public-release data. In addition, both EHC and Q-list instruments collected information on respondents’

children, on their educational histories, and on their parental upbringing. Figure 6 provides a screen shot of the life course EHC that was used.

In analyses, the retrospective responses were compared with those from the same respondents collected over the history of their participation in the PSID for purposes of validation. Complete analyses have been conducted for the social and economic outcome variables, including residential changes, and marriage, cohabitation, employment, and unemployment histories. These results have been presented at national and international conferences (Andreski, Belli, & Stafford, 2003; Belli, 2004; Belli, 2003; Belli, Andreski, & Agrawal, 2004). Analyses have also been completed for cigarette smoking, and initial analyses have been conducted on reports of health status.

As predicted, for most outcome measures, the EHC outperformed the Q-list in data quality, including the number of residential changes, the number of years and which years respondents were engaged in cohabitation relationships, the amount of annual employment and which years employment took place, the number of years unemployed and which years respondents were unemployed, and the amount of cigarette smoking. There were no differences between conditions in reports of which years residential changes were made. Uncharacteristically, retrospective marriage reports were found to correspond more strongly on the number of marriages and the date of first marriage with the panel data in the Q-list interviews. This latter result, however, is likely due to the failure in the EHC interviews to ask for the number of times married as was done in the Q-list interviews; instead respondents were asked in EHC interviews for the first names or initials of all intimate partners with which respondents had lived. Apparently, how many times one is married is information that is pre-stored in memory, and having reported this number assists in remembering specific marriages. On the other hand, respondents likely had forgotten some married partners in the EHC condition when being asked to provide names or initials of partners. In EHC interviewing then, it is advisable to ask respondents how many times that they had been married (see Freedman et al., 1988).

Because of space limitations, not all analyses will be reported. Figures 6 and 7 provide an illustration of two outcome measures and the differences in correspondence with the panel data between EHC and Q-list conditions. In Figure 7, annual correlations

between retrospective amounts of work and panel work hours are charted, and in Figure 8, annual kappa indices for retrospective and panel reports of cohabitation. In both cases, the EHC provides better correspondence with the panel reports than the Q-list overall ( $\chi^2(1) = 6.13, p = .01$  for amount worked,  $\chi^2(1) = 4.65, p = .03$  for cohabitation). However, the patterns across time in both of these figures are markedly different. In the reports of annual amounts of work, the EHC is noted for producing better quality data for the remote years, whereas the Q-list gains on the EHC for the more recent years. As for reports of cohabitation, the EHC is noted for markedly superior correspondence in comparison to the Q-list for the more recent years. These differences may partly reflect the chronological order of retrieval that was scripted in the Q-list: a backward order of retrieval for employment history and a forward order of retrieval for cohabitants. Importantly, a verbal behavior analyses that examines order of retrieval could not only support or refute these speculations, but can also reveal how order of retrieval affects data quality in the EHC condition as well, in which interviewers and respondents negotiated for each domain a forward, backward, or mixed order of retrieval.

The retrospective reports of smoking also indicate that the EHC condition provides higher quality data. The PSID has limited panel data on whether respondents currently smoke, and the number of cigarettes smoked, having asked these questions only in 1986, 1999, and 2001. These dates correspond to retrospective reports of events that had occurred approximately 16, 3, and 1 year previously. Both the EHC and Q-list conditions collected annual life course reports using 4 categories of *not having smoked at all*, or of having smoked on average between *1 and 10*, *11-20*, or *21 or more* cigarettes per day. For analytic purposes, kappa indices were computed between retrospective and panel reports to measure agreement using all 4 categories, using 3 categories by collapsing the middle categories resulting in a middle category of smoking 1-20 cigarettes, and using two categories of not smoking at all, and smoking 1 or more cigarettes per day. These analyses are reported in Table 2, which reveals that the levels of agreement as measured by kappa improve for more recent years (consistent in the EHC condition), and improve for measures that included fewer categories. Most striking is that the EHC condition had stronger kappa levels in comparison to the Q-list in all comparisons, with about half leading to a statistically significant difference.

Another variable, reports of health status, have also been analyzed at the time of this writing. Both panel and retrospective reports contain annual reports of health status on a scale of 1 = excellent to 5 = poor. The panel data are constrained in that health status was not asked until 1984; thus analyses are based on reports for 1984 through 1997. For both EHC and Q-list conditions, intraclass correlations were consistently moderately strong as they ranged from .45 to .65. Although the EHC condition had more years ( $n = 9$ ) in which there were higher intraclass correlations than the Q-list condition ( $n=5$ ), a goodness of fit test did not yield a significant difference.

A series of mixed-model analyses were conducted to determine whether there were slope differences in the annual mean levels of health status between retrospective and panel (prospective) reports between EHC and Q-list conditions. The mean levels of health status across years are provided in Figure 9. Most noteworthy, the slopes for all observations are positive and significantly different from 0, indicating that respondents' were reporting worse health status as they aged both prospectively, and retrospectively. Interestingly, the slope of the Q-list retrospective data (slope = .015) is significantly shallower than the corresponding prospective data (slope = .024,  $t(612) = 2.08$ ,  $p = .04$ ). The slope of the EHC retrospective reports (slope = .021) did not significantly differ from the corresponding prospective data (slope = .022).

Again using a mixed-model approach, the predicted mean levels were examined to determine the presence of significant differences between retrospective and panel (prospective) reports at the beginning year of the series (1984), at the midpoint (between 1990 and 1991), and at the ending year of the series (1997). For 1984, it was found that only the EHC condition had a lower predicted mean retrospectively in comparison to the corresponding panel predicted mean,  $t(617) = 2.23$ ,  $p = .03$ . However, for both the midpoint, and for 1991, both EHC and Q-list conditions yielded predicted mean values that were significantly lower than those observed in their respective corresponding panel data (midpoint Q-list  $t(618) = 2.42$ ,  $p = .02$ ; 1997 Q-list  $t(615) = 3.32$ ,  $p = .001$ ; midpoint EHC  $t(617) = 2.81$ ,  $p < .01$ ; 1997 EHC  $t(612) = 2.63$ ,  $p < .01$ ).

In summary, respondents in both EHC and Q-list conditions were retrospectively reporting their health status as better in comparison to their prospective panel reports. Only the EHC, however, preserved the rate of change over time that was observed in the

panel data. As for the Q-list, respondents reported less deterioration in health status over time retrospectively than observed in the corresponding panel reports.

In addition to these data quality measures, operational measures including interviewing time and interviewer assessments were analyzed. Although the EHC condition did lead to significantly longer interviews ( $M = 57.6$  minutes,  $SD = 28.7$ ) in comparison to the Q-list condition ( $M = 51.5$ ,  $SD = 21.9$ ),  $t(615) = 2.96$ ,  $p = .003$ , the 6-minute average difference between conditions is relatively modest in the context of a nearly hour-long interview. Noteworthy is that interviewers preferred the EHC. They found respondents to be significantly more cooperative and motivated in the EHC condition in comparison to the Q-list. In addition, interviewers perceived respondents to be better able to understand and answer questions, and to remember both major and detailed past events, in the EHC condition. Finally, interviewers reported to have enjoyed the EHC interviews more, they found EHC interviews as being easier to administer and to record answers, and that the EHC computer application facilitated interviewing more than the Q-list application. These results largely replicate Belli et al.'s (2001) findings that interviewers preferred paper and pencil EHC interviews over Q-list ones.

### Conclusion

Methodological research has shown that the Event History Calendar interviewing methodology leads to consistently higher quality retrospective reports in comparison to traditional standardized Q-list methods. These improvements to data quality have been observed in both paper and pencil and computerized instruments, and with 2-year and life course reference periods. Costs in interviewing time are negligible or nonexistent. The 2003 PSID illustrates that computerized EHC interviewing can be successfully implemented in core panel interviews. In conclusion, following the lead of the PSID, social and health scientists should seriously consider benefits that calendar-based interviews can provide to the quality of retrospective survey data to which they oft rely upon for their scientific inferences.

## References

- Andreski, P., Belli, R. F., & Stafford, F. P. (2003, July). Computerized calendar and question list methods: Economic and health measures Poster presented at the biennial meeting of the Society for Applied Research in Memory and Cognition, Aberdeen, Scotland.
- Axinn, W. G., Barber, J. S., & Ghimire, D. J.. (1997). The neighborhood history calendar: A data collection method designed for dynamic multilevel modeling. In A. E. Raftery (ed.), *Sociological Methodology* (Volume 27, pp. 355-392). Washington, DC: American Sociological Association.
- 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, 383-406.
- Belli, R. F. (2000). Computerized event history calendar methods: Facilitating autobiographical recall. *American Statistical Association Proceedings of the Section on Survey Research Methods* (pp.471-475). Alexandria, VA: American Statistical Association.
- Belli, R. F. (2004, August). Improving the quality of retrospective reports: Calendar interviewing methodologies. Paper presented at the sixth international conference on logic and methodology, Amsterdam, the Netherlands.
- Belli, R. F. (2003, November). Improving autobiographical memory recall in surveys: Calendar interviewing methodologies. Paper presented at the annual meeting of the Psychonomic Society, Vancouver, British Columbia.
- Belli, R. F., Andreski, P., & Agrawal, S. (2004, May). CATI event history calendar and question list methods: Accuracy of life course retrospective reports. Paper presented at the annual meeting of the American Association for Public Opinion Research, Phoenix
- Belli, R.F., Shay, W. L., & Stafford, F. P. (2001). Event history calendars and question list surveys: A direct comparison of interviewing methods. *Public Opinion Quarterly*, 65, 45-74.
- Belli, R. F., Lee, E. H., Stafford, F. P., & Chou, C-H. (2004). Calendar and question-list survey methods: Association between interviewer behaviors and data quality. *Journal of Official Statistics*, 20, 185-218.
- Brown, N. R., & Schopflocher, D. (1999). Event clusters: An organization of personal events in autobiographical memory. *Psychological Science*, 9, 470-475.
- Cannell, C. F., Fisher, G. & Bakker, T. (1965). Reporting of hospitalization in the Health Interview Survey. *Vital and Health Statistics* (PHS Publication No. 1000, Series 2, No. 6). Washington, DC: U.S. Government Printing Office.
- Caspi, A., Moffitt, T. E., Thornton, A., Freedman, D., Amell, J. W., Harrington, H., Smeijers, J., & Silva, P. A. (1996). The life history calendar: A research and clinical assessment method for collecting retrospective event-history data. *International Journal of Methods in Psychiatric Research*, 6, 101-114.
- Dugoni, B., Lee, L., & Tourangeau, R. (1996). Report on the NLS/Y round 16 recall experiment. Unpublished manuscript: National Opinion Research Center.

- Duncan, G. J., & Hill, D. H. (1985). An investigation of the extent and consequences of measurement error in labor economic survey data. *Journal of Labor Economics*, 3, 508-522.
- Freedman, D., Thornton, A., Camburn, D., Alwin, D., & Young-DeMarco, L. (1988). The life history calendar: A technique for collecting retrospective data. In C.C. Clogg (ed.), *Sociological Methodology* (Vol. 18, pp. 37-68). San Francisco; Jossey-Bass.
- Furstenberg, F. F. Jr., Brooks-Gunn, J., & Morgan, S. P. (1987). *Adolescent mothers in later life*. New York: Cambridge University Press.
- Kessler, R. C., & Wethington, E. (1991). The reliability of life event reports in a community survey. *Psychological Medicine* 21, 723-738.
- Linton, M. (1982). Transformations of memory in everyday life." In U. Neisser (ed.), *Memory observed: Remembering in natural contexts* (pp. 77-91). San Francisco: Freeman.
- Lyketsos, C. G., Nestadt, G., Cwi, J., Heithoff, K., & Eaton, W. E. (1994). The life chart interview: A standardized method to describe the course of psychopathology. *International Journal of Methods in Psychiatric Research*, 4, 143-155.
- Mathiowetz, N. A., & Duncan, G. J. (1988). Out of work, out of mind: Response errors in retrospective reports of unemployment. *Journal of Business and Economic Statistics*, 6, 221-229.
- Rubin, D. C., & Wenzel, A. E. (1996). "One hundred years of forgetting: A quantitative description of retention. *Psychological Review*, 103, 734-760.
- Schank, R. C., & Abelson, R. P. (1995). Knowledge and memory: The real story. In R. S. Wyer (ed.), *Advances in social cognition* (Vol VII, pp. 1-85). Hillsdale, NJ: Erlbaum.
- Thompson, C. P., Skowronski, J. J., Larsen, S. F., & Betz, A. L. (1996). *Autobiographical memory*. Mahway, NJ: Erlbaum
- Van der Vaart, W. (2004). The time-line as a device to enhance recall in standardized research interviews: A split ballot study. *Journal of Official Statistics*, 20, 301-318.
- Wagenaar, W. A. (1986). My memory: A study of autobiographical memory over six years. *Cognitive Psychology* 18, 225-252.



Figure 1.

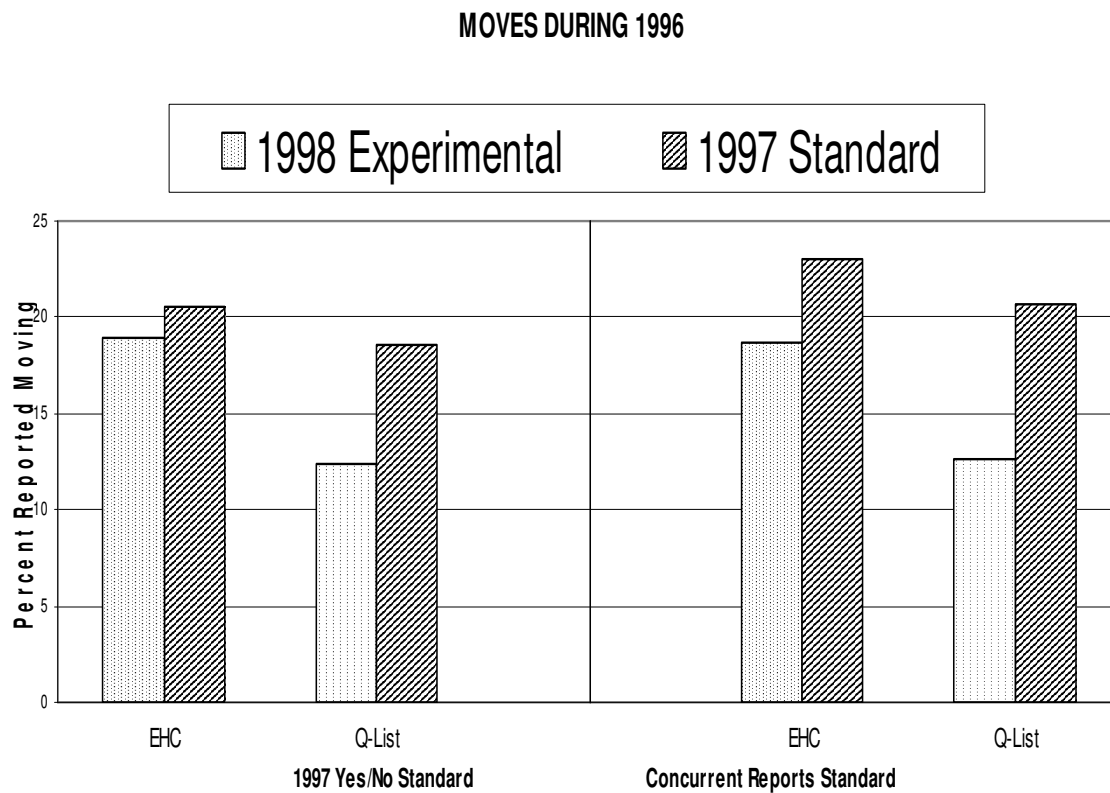


Figure 2.

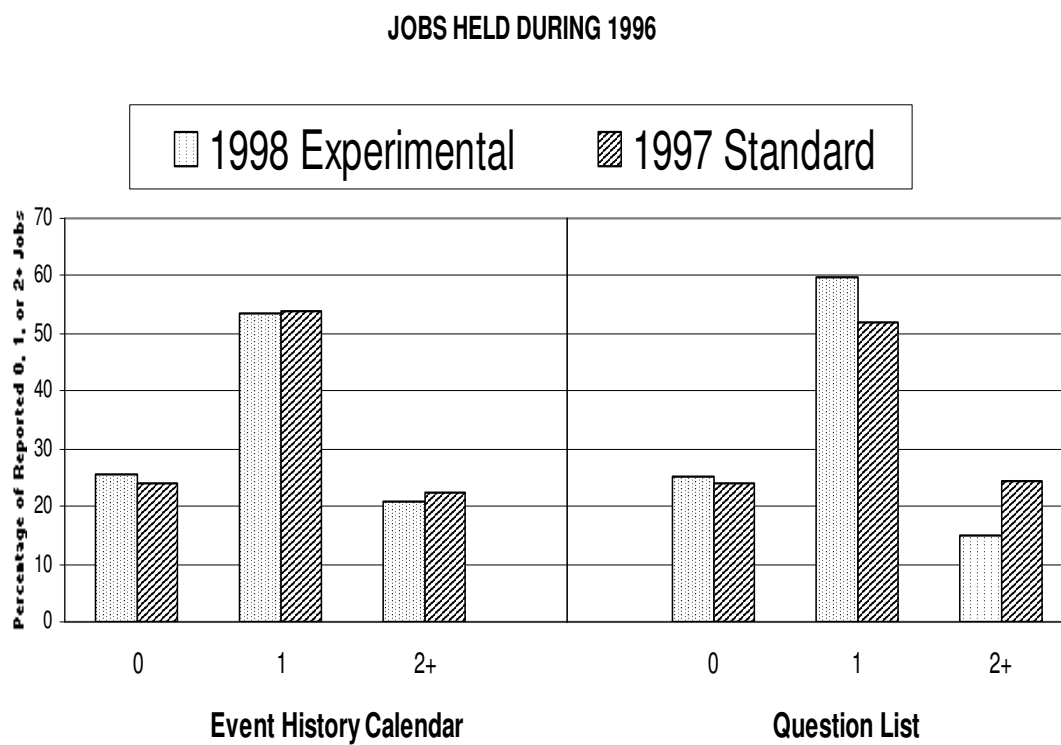


Figure 3.

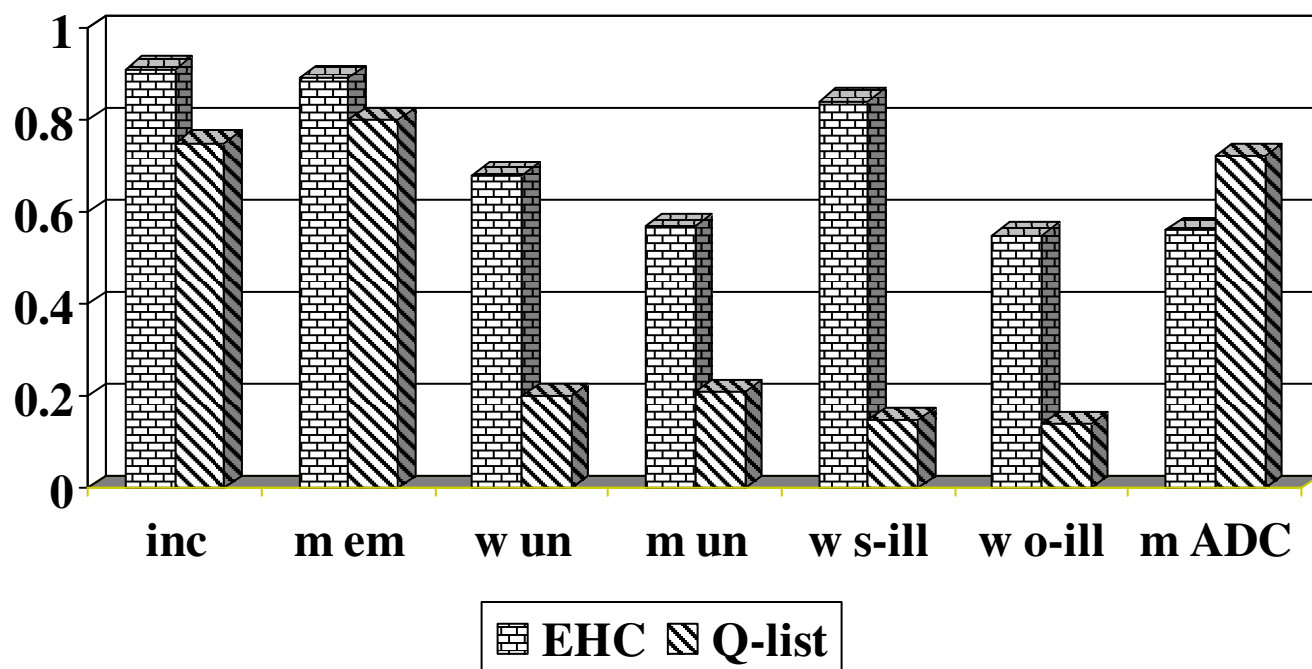


Figure 4

The screenshot shows the 'Event History Calendar - [Employment (About) Phil]' application window. The interface is divided into three main sections:

- Top Section:** Contains a calendar grid for the year 2002, with months from SPR to WIN. Below the grid are sections for 'Landmark Events', 'Residence', 'Employment Summary', and 'Not Working'. A tooltip is visible over the calendar, stating: '321 Real Nice Place, Ann Arbor, MI [From 09/11/2002 to 02/21/2003]'. A tabbed interface below shows 'Landmark Events', 'Residence', 'Employment', 'Not Working', and 'TimeAway', with 'Employment' selected.
- Middle Section:** Titled 'Employment Data Entry Window', it features a 'Not Working Summary' and a grid for entering data for various employers: 'Joe and Sons', 'Bill's Landscaping', 'Job# 3', and 'Job# 4'. Each employer has a corresponding row of colored blocks representing work periods.
- Bottom Section:** Contains a form for entering specific data. It includes 'Starting Time' and 'Ending Time' fields, each with dropdowns for Year, Season, Month, and 3rd of Month. The 'Employer' field is set to 'Bill's Landscaping'. There is a 'Note' text area and buttons for 'OK', 'Cancel', and 'Delete'.

The Windows taskbar at the bottom shows the Start button, open applications including 'C:\WINDOWS\Des...', 'Event History Ca...', and 'Microsoft PowerPain...', and the system clock showing 3:33 PM.

Figure 5

**Event History Calendar - [Employment]**

File Window Help

WIN 01	SPR 01	SPR 01	SUM 01	FAL 01	FAL 01	WIN 02	SPR 02	SPR 02	SUM 02	FAL 02	FAL 02
JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC

**Landmark Events**

**Residence**

**Employment Summary**

**Not Working**

---

**Employment Q-list**

**Q1. We would like to know about what you do**  
 Are you working now, Are you looking for work, Are you retired, Are you keeping house, Are you a student, or what?

Working now  
 Only temporarily laid off, sick leave, or maternity leave  
 Looking for work, unemployed  
 Retired In what year did 'you' retire? 1995  
 Disabled permanently or temporarily  
 Keeping house  
 Student  
 Other  
 DK  RF

**Q2. Have you done any work for money since January 1, 2001 ? Please include any type of work, no matter how small.**

Yes  No  DK  RF

Q-list (1, 2)      Q-list (3,4,5)      Employment

Table 1: Correspondence of Data between 2001 and 2003 PSID

2001 Q-list Questionnaire		2003 (EHC) Questionnaire
<u>Underlining indicates differences in the data collected between years</u> <i>Italics indicate questions that EHC does not include</i> <b>Bold indicates additional information collected by EHC</b>		
Questions/Variables	Data description	Questions, availability
A42 – A47 ER17088-ER17090	Whether moved since last interview, (if yes) month and year of most recent move, why move, might move, likelihood of move, why might move	A49-A53 Available, <b>thirds-of-a-month of all moves (if any) since January 2001 until interview date</b>
B1-B3 ER17216–ER17220	up to 3 mentions of employment status, year retired, <i>work for money now</i>	BC1-BC3 Mostly available; <i>work for money now not available in EHC</i>
B4 – B19 ER17221- ER17244	Current main job: other-or-self employed govt or private company union contract belong to labor union occupation/ activities/ industry method of payment (hourly, salary) and amount overtime pay	BC4-BC6a; BC15-BC37; All available; <b>also have for current year have whether worked entire month, part of month, or not at all for each job interval of continuous work (job spell); have start and stop dates to thirds-of-a-month.</b>
B23 ER17254-ER17256	Years, months, weeks experience with current employer	BC41 All available
B24 ER17257-ER17258	Month, (day), and year of starting work with current employer	Available, <i>but default value in EHC program led to many responses of January 1, 2001 – this data collection problem will be corrected for 2005.</i>
<i>If working now:</i> B25 – B38 ER17259-ER17275	<i>Current employer: Month and year of starting position Position changes in last calendar year (T-1), and whether higher pay occupation/ activities/ industry when starting work in last calendar year (T-1) starting salary when starting in last calendar year (T-1) hours per week when started</i>	<i>Not available</i>

<p>B39 – B45B ER17276-ER17308</p>	<p>Months in last calendar year (T-1) working for employer for current job Any (other) main-job employer during last calendar year (T-1) occupation/ activities/ industry Months in last calendar year (T-1) working for that employer Whether overlap in weeks in working for other main job other-or-self employed govt or private company How much made in dollars in last calendar year (T-1) Hours per week worked in last calendar year (T-1)</p>	<p>BC4-BC6a; BC15-BC22; BC42-BC46; BC53-BC56 <b>For months in last calendar year (T-1), have whether worked entire month, part of month, or not at all for each job interval of continuous work (job spell); have start and stop dates to thirds-of-a-month. main job(s) overlap indicator can be more precise to not include months in which the same third-of-months do not overlap</b></p>
<p><i>B46 – B52 ER17309-ER17317</i></p>	<p><i>Position changes in last calendar year (T-1), and whether higher pay Month, (day), and year of starting work If starting before last calendar year (T-1), occupation and activities, starting salary and hours per week</i></p>	<p><i>Not available</i></p>
<p>B53-B55 ER17318-ER17321</p>	<p>Whether stopped working for employer of other main job held in current year or last calendar year (T-1) Month and year stopped working Reason for stopping work Note: corresponds in C section for persons who are not currently working</p>	<p>BC4-BC6a: BC47-BC51 Available; <b>timing of stopping work to third of month</b></p>
<p><i>B56-B57a ER17318-ER17321</i></p>	<p><i>Amount making when stopping work Hours per week when stopping work</i></p>	<p><i>Not directly available</i></p>

<p>B59  <u>No data center variable</u>  <u>No B58 question #</u></p>	<p>Any other main jobs in current and last calendar years  Possible to locate how many different jobs worked in current and last calendar year (T-1), but not in data center</p>	<p>BC4-BC6a; BC52  EHC permits the ability to assess main jobs and other jobs for employment in interview year and for employment that occurred in last calendar year (T-1); specifically can assess how many different jobs worked, the number of weeks worked for each job, in which months (<b>and thirds-of-month</b>) one was working at each job, and the average hours per week worked per job.</p>
<p>B79  ER17393-ER17394</p>	<p>On average, hours per week on all <b>main jobs</b> in last calendar year (T-1)</p>	<p>BC4-BC6a; BC43  <b>number of weeks worked and hours per week available for all job spells, not just main jobs, but problem with intermittent work</b></p>
<p>B80 – B81d  ER17391-ER17393</p>	<p>Whether overtime not included in hours per week on main jobs  Total overtime hours in last calendar year (T-1)  Separate answers on hours per week and total overtime hours for each main job</p>	<p>All available</p>



<p>B82 – B106 ER17398-ER17502</p>	<p>Extra jobs in last calendar year (T-1) that were not main jobs govt or private company occupation/ activities/ industry How much made in dollars in last calendar year (T-1) Weeks worked in last calendar year (T-1) Hours per week worked in last calendar year (T-1) Month and year started working Which months in last calendar year (T-1) working Whether stopped working for employer Month and year stopped working Cycles through for all extra jobs in last calendar year (T-1)</p>	<p>BC4-BC6a; BC15-BC22; BC42-BC46; BC52 All available, <u>but notion of “extra job” not included in calendar, only main jobs and other jobs can be extracted</u></p>
<p>B20 – B21 ER17245-ER17253</p>	<p>If working now: Looking for another job? If yes, what has been done to find one For those not working now, how long been looking for work asked in section C</p>	<p>BC64-BC67 All available</p>

<p>B60 – B78 ER17325-17392</p>	<p>Number of weeks unemployed, out-of-labor force, missing work, and actually working in last calendar year (T-1) Month strings for unemployment and out-of-labor Available only for last calendar year (T-1)</p>	<p>BC7-BC14 All available; <b>For month strings, can include whether status persisted for entire month, or part of month.</b> <b>Calendar separates out unemployment and temporary lay-off, Q-list has them both combined.</b> <b>For unemployment, can determine how many different spells of unemployment occurred</b></p>
<p>R25 – R45 These variables are located in T-2 Individual Income Files Supplemental Data</p>	<p>For head/wife Whether job/business two calendar years ago (T-2) How much earned two calendar years ago (T-2) How many weeks worked Which months worked in two calendar years ago (T-2) Hours per week worked in months worked two calendar years ago (T-2) Which months unemployed in two calendar years ago (T-2)</p>	<p>BC4-BC9; BC47-BC52; BC57-BC60 All available; <b>also for months in two calendar years ago (T-2), have whether worked entire month, part of month, or not at all for each job interval of continuous work (job spell); parallel information for unemployment; parallel information for out-of-labor force; have start and stop dates to thirds-of-a-month.</b> <b>Hours per week and how much made per job spell available for employment that occurred two calendar years ago (T-2).</b> <b>Whether stopped work, and reasons for stopping work available for T-2.</b> <b>Can determine how many different jobs held in T-2.</b> <b>For unemployment, can determine how many different spells of unemployment occurred</b></p>

<p>Work Hours and Wages – last calendar year (T-1) only variables; All variables are relevant</p>	<p>Work weeks for head and wife, work hours for head and wife, overtime hours for head and wife, work hours on extra jobs for head and wife, total hours of work for head and wife, wage rate for head and wife, weeks missed due to others' illness for head and wife, weeks missed due to self-illness for head and wife, vacation weeks for head and wife, weeks on strike for head and wife, unemployment weeks for head and wife, weeks out-of-the labor force for head and wife</p>	<p>Available (note exception below)  <b>Work weeks and work hours for head and wife also available for T-2</b>  <b>Unemployment weeks for head and wife also available for T-2, but if adding temporarily laid off, only available for T-1</b>  <b>Out-of-the labor force weeks for head and wife also available for T-2</b>  <i>Not clear how to extract work hours on extra jobs, can have work hours on jobs other than main jobs</i></p>
---	---	--

Figure 6. Screen capture of the lifecourse EHC for the labor history domain. The EHC is copyright (c) 2000 to the Regents of the University of Michigan.

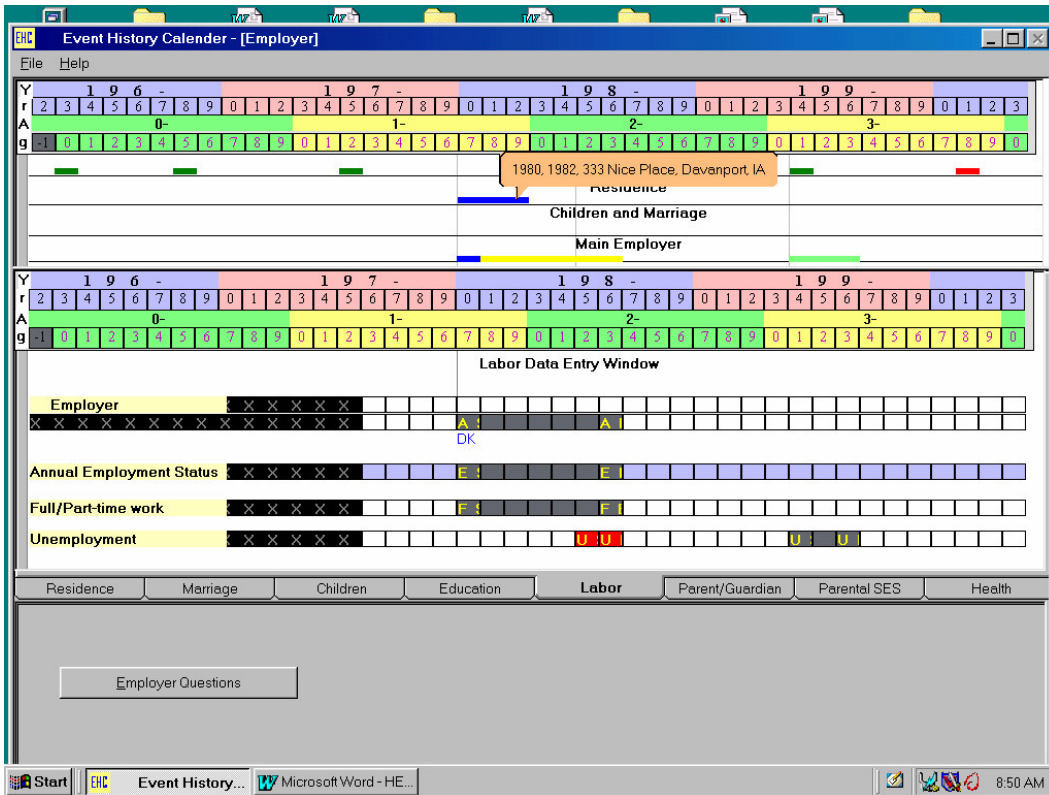


Figure 7: Correlations between annual retrospective amounts worked and panel work hours

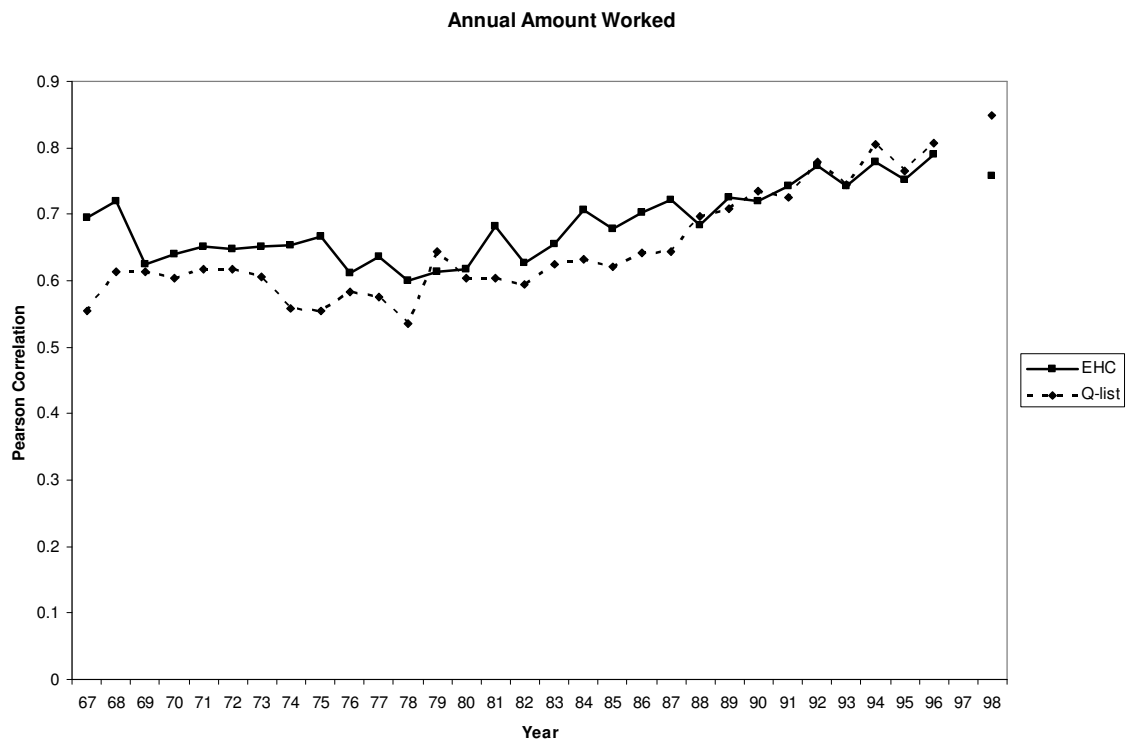


Figure 8: Kappas between annual retrospective and panel cohabitation partners

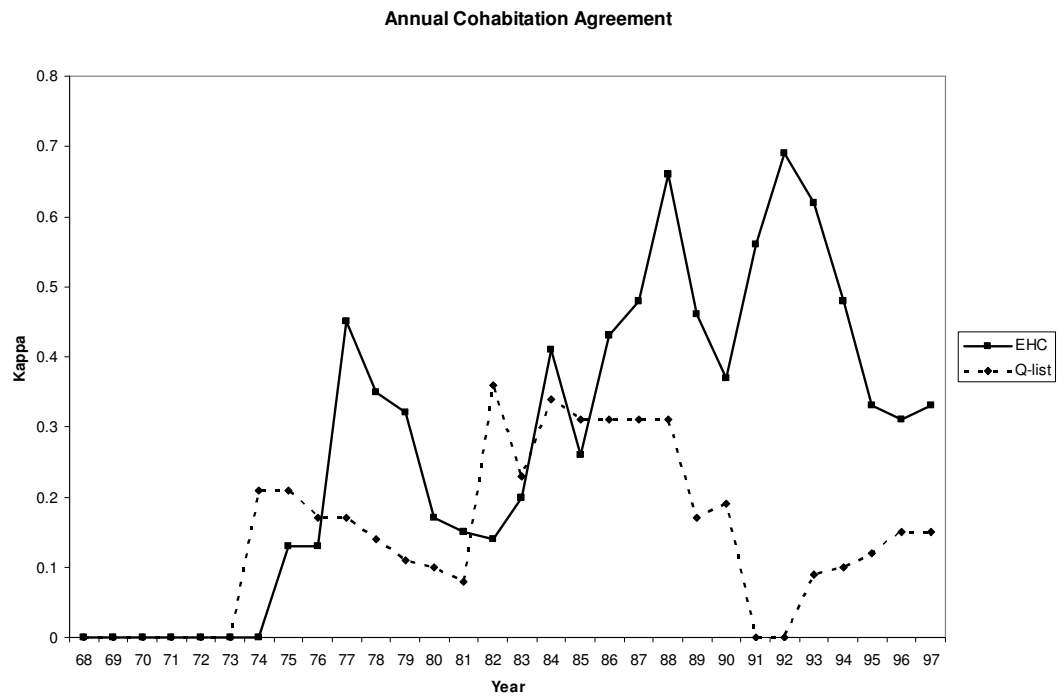


Table 2. Kappa indices on amount smoking for EHC and Q-list conditions. Bold indicates indices that are significantly different between conditions.

Amount smoked	1986			1999			2001		
	EHC	Q-list	p	EHC	Q-list	p	EHC	Q-list	p
4 categories	<b>.59</b>	<b>.48</b>	<b>.04</b>	.72	.62	.11	<b>.73</b>	<b>.48</b>	<b>&lt;.01</b>
3 categories	.64	.60	.54	.78	.67	.09	<b>.79</b>	<b>.61</b>	<b>&lt;.01</b>
2 categories	.79	.74	.40	.84	.75	.12	<b>.85</b>	<b>.73</b>	<b>.04</b>

Figure 9. Mean health status for panel (prospective) and retrospective reports.

