Justice Quarterly

ISSN: 0741-8825 (Print) 1745-9109 (Online) Journal homepage: https://www.tandfonline.com/loi/rjqy20

Assessing the Accuracy of Life Event Calendar Data in an Offender Sample

Nikki van Gerwen, Arjan Blokland & Arieke J. Rijken

To cite this article: Nikki van Gerwen, Arjan Blokland & Arieke J. Rijken (2019) Assessing the Accuracy of Life Event Calendar Data in an Offender Sample, Justice Quarterly, 36:3, 532-566, DOI: 10.1080/07418825.2018.1458143

To link to this article: https://doi.org/10.1080/07418825.2018.1458143

Published online: 09 Apr 2018.

Justice Quarterly
2019, Vol. 36, No. 3, 532–566
https://doi.org/10.1080/07418825.2018.1458143

Assessing the Accuracy of Life Event Calendar Data in an Offender Sample

Nikki van Gerwen , Arjan Blokland and Arieke J. Rijken

Criminologists are increasingly interested in the effects of life-course dynamics on criminological development. However, detailed longitudinal data are difficult to obtain and possibly confounded due to recall errors. Life Event Calendars (LECs) are designed to reduce recall errors and are increasingly used as a method for obtaining valid retrospective data in criminological studies. Yet few studies exist that assess the accuracy of LEC data in offender samples. This study aims to fill this void. We compare data regarding the prevalence and timing of marriage, divorce, and childbirth obtained through an LEC to official registry data in a sample of convicted offenders. We examine whether the accuracy of the data vary by event or respondent specific characteristics. We conclude that the LEC data are quite accurate regarding the prevalence of marriage, divorce, and childbirth. The data are less accurate regarding the timing of these life events.

Keywords Life Event Calendar; data accuracy; validity; self-reported data; offender sample

Nikki van Gerwen is a PhD candidate at the Department of Sociology at Utrecht University/ICS. Her research interests are in the area of cooperation, social dilemmas, cohesion and experimental research. Arjan Blokland is a senior researcher at the Netherlands Institute for the Study of Crime and Law Enforcement, Amsterdam and professor of Criminology and Criminal Justice at Leiden Law School, Leiden University, The Netherlands. His research interests include developmental and life-course criminology, effects of criminal justice interventions, and sexual violence. Arieke J. Rijken is an assistant professor at the Department of Sociology, Vrije Universiteit Amsterdam. Her research interests are in the areas of life-course sociology, life-course criminology and political participation. Correspondence to: Nikki van Gerwen, Department of Sociology, Utrecht University, Padualaan 14, 3584 CH Utrecht, The Netherlands. E-mail: N.m.h.vangerwen@uu.nl
This article was originally published with errors. This version has been corrected. Please see Corrigendum (https://doi.org/10.1080/07418825.2018.1474562).

Introduction

Inspired by insights from developmental psychology and life-course sociology, many criminologists have come to embrace a life-course approach to understanding crime and deviance (Blokland & Nieuwbeerta, 2010). With the advent of life-course criminology, the field has witnessed increasing attention for within-individual variability in criminal behavior over longer periods of time. This heightened interest in life-course dynamics has in turn increased the demand for detailed longitudinal data (Caspi et al., 1996). Longitudinal research designs are needed to accurately disentangle processes related to criminal behavior over the life-course (Farrington, 1988). Cross-sectional designs preclude separating the effects of extrinsic variables from the effects of enduring individual differences. To take the study of marriage and crime as a well-known example: to find that married men commit less crime than unmarried men allows for less stronger causal inferences, than to find that these same unmarried men commit less crime during the years they are in a marital relationship. Rather than differences in criminal propensity between married and unmarried men, the latter finding would strongly suggest a beneficial effect of social bonds on crime, which in turn would favor policies aimed at increasing offenders' social capital to reduce crime (see Liberman, 2008:8).

Longitudinal data are thus needed to precisely map if and when important events and transitions, such as marriage, occur during the course of individuals' lives. Ideally, researchers use well-designed prospective observational studies for this purpose. Two common problems with such designs are the large costs and long duration until the data are available for analysis. These problems can be avoided by using retrospective-survey designs, such as Life Event Calendars (LECs) (Morris & Slocum, 2010).

LECs are used as a data collection method for obtaining reliable retrospective data. The LEC was developed in the context of longitudinal research to obtain information about central events that can occur in a respondent's life (Caspi et al., 1996). A possible complication in retrospective studies is the recall bias that may lead to under or over reporting and inaccurate timing of life events, as reported by respondents (Eisenhower, Mathiowetz, & Morgenstein, 1991). Recall errors in retrospective surveys may affect both recall and dating accuracy of respondents regarding specific life events, which in turn may lead to erroneous conclusions about the effects of life-course transitions on other life trajectories, such as criminal development. The LEC tries to minimize these recall errors by using visual aids, inquiring about streams of events, and contextualizing questions about various life events (Brener, Billy, & Grady, 2003). It enhances recall through processes that capitalize on the sequential and hierarchical storage of memories (Belli, 1988; Morris & Slocum, 2010). Especially in criminological applications, the LEC is increasingly used as a stimulation tool for the memory when collecting time-based data (Sutton, 2010).

The use of self-report measures invokes questions regarding the reliability and validity of the data obtained (Auriat, 1993; Elliott, 1982). Extensive research has been conducted on the reliability and validity of self-report measures that capture the prevalence of self-reported criminal events (e.g. Hindelang, Hirschi, & Weis, 1981; Thornberry & Krohn, 2000). The move towards the life-course approach has caused more recent work to rely on retrospective longitudinal data with regard to events and transitions in other life-course domains as well (e.g. Slocum, Simpson, & Smith, 2005; Uggen & Thompson, 2003). Methodological assessments of the reliability and validity of self-report data have failed to keep pace with these theoretical advances. The shift towards the life-course approach furthermore requires reliability and validity assessments not only on the *prevalence* of specific life events, i.e. the number of times a life event occurs, but also on the *timing* of these specific life events. Little is known about the LEC's ability to elicit accurate self-reports on the prevalence and timing of specific life events, especially over the entire life-course among offenders, the main population of interest for life-course criminologists. The consequences of this oversight are potentially severe, considering that dating accuracy problems can fundamentally compromise the quality of self-reported data (Belli, 1988).

The present study aims to fill this void by examining the accuracy of self-reported data, collected using LEC-methods in an offender sample. Specifically, we pose two questions: (1) *How accurate are self-reported data on the prevalence and timing of marriage, divorce, and childbirth collected using Life Event Calendar methods in an offender sample?* and (2): *Does the accuracy of these data vary systematically by characteristics of the respondent or the event?* We focus on marriage, divorce, and childbirth since we have data on these life-course domains from an external register and data gathered through an LEC available. Examining the concordance between these two datasets allows us to draw conclusions on the accuracy of the information obtained through LECs.

The scarce research that has assessed the accuracy of LEC data collected among offender samples is limited to American prisoner samples and focused on validating data on arrests (Morris & Slocum, 2010; Roberts & Wells, 2010) and jail and prison terms (Roberts & Wells, 2010; Sutton, Bellair, Kowalski, Light, & Hutcherson, 2011). These studies did not compare self-reported data on events and transitions in other life-course domains to official data. Although concordance between self-reported data and official data on (lifelong) number of jail and prison terms was fairly high, these results do not bear directly on the accuracy of the LEC because the number of jail and prison terms were asked with "standard question" and did relate to timing, which is the most important aspect of the LEC (Roberts & Wells, 2010; Sutton et al., 2011).

Rather, these results address the validity of prisoners self-reports in general. Frequency and especially timing of arrests have proven to be more difficult to recall accurately. Roberts and Wells (2010) as well as Morris and Slocum (2010) compared monthly LEC data, each study covering a three-year period, to official data on arrests and reported high levels of recall error.[1]

The contribution of the current study is primarily methodological. We aim to assess the accuracy of retrospective data by comparing data obtained through LECs with data obtained from an external dataset, in this case data obtained through centralized municipal registrations. This study contributes to the existing literature in at least two ways: first, it is one of the few to examine the concordance between data from LECs and external data sources over life-long recall periods. Second, in doing so, this study makes use of a sample of offenders, a highly interesting subpopulation. Collecting time-based data of offenders is especially challenging because their "lives often are chaotic and unstable, with frequent changes in residence, education, employment, and interpersonal relationships" (Schubert et al., 2004, p. 7). To our knowledge, this study is the first to combine these two elements, which allows us to draw conclusions on the accuracy of the use of LEC data for criminological research.

In answering the posed research questions, this study makes use of data from the Criminal Career and Life-Course Study (CCLS) (Blokland, 2005). The CCLS study gathered a large-scale longitudinal dataset with both self-reported, as well as official data on conviction histories and the life-events of our interest in a Dutch offender sample. The data also include a non-offender reference group of which we take advantage in this study.

Literature Review and Hypotheses

The Life Event Calendar: A Brief Introduction

Over the past decades, the LEC method has emerged as an extension of traditional self-report surveys. The fundamental idea of life-course theory, that individual lives can be understood as a set of experiences and events that are interconnected and mutually reinforcing (Elder, 1985; Wheaton & Gotlib, 1997), has been a guiding principle in the development of the LEC (Freedman, Thornton, Camburn, Alwin, & Young-DeMarco, 1988). As the name implies, the LEC is structured in a calendar format with predetermined time periods positioned horizontally and life domains of interest positioned vertically (Morris & Slocum, 2010). The LEC is usually structured around three main elements: the recall period, a time unit, and the life domains. Interviewers use LECs to collect informa-

1. Percentages of errors were especially high when person-months without arrests (according to the official data) were excluded from the analysis. This is, in our view, the best way to look at the accuracy of self-reported arrests, as the high percentage of months without arrests will "artificially" increase the concordance between self-reported and official data.

tion on a respondent's life events that occurred within a designated period of time. During the interview, interviewers and respondents typically work together to chart when various life events occurred in the respondent's life. The main advantage of using LECs as data collection method is that it is designed to reduce recall errors (Belli, 1988; Caspi et al., 1996; Morris & Slocum, 2010; Stafford, 2009).

The mode of administration adopted by LECs enables researchers to efficiently collect complicated longitudinal data from respondents (Axinn, Pearce, & Ghimire, 1999). LECs are more interactive than traditional retrospective self-report surveys as the interviewer works together with the respondent to fill out the calendar, beginning with the events most easily remembered and working from there to fill out the rest of the calendar (Bellie, 2000). This dynamic approach allows interviewers and respondents to locate possible inaccuracies and adjust these during the interview. The LEC presents respondents with both mental (questions) and visual (the calendar) aids to facilitate recall (Axinn et al., 1999; Laub & Sampson, 2003). In doing so, this methods taps into more competences than traditional self-reports and draws from a larger pool of memories, thereby reducing recall errors (Horney, 2001). Experimental studies have found that the LEC methodology produced improvements to recall compared to traditional survey approaches (Belli, Shay, & Stafford, 2001; Belli, Smith, Andreski, & Agrawal, 2007).

The LEC method may be particularly useful for gathering data from populations whose members feature chaotic lives, reading difficulties, or memory problems, and who experience multiple changes over their life-course (Engel, Keifer, & Zahm, 2001). The LEC method has been incorporated widely in criminological research (Roberts & Horney, 2009) and has been used in studies focusing on life-course trajectories of juvenile delinquents (Laub & Sampson, 2003), victimization (Wittebrood & Nieuwbeerta, 2000), criminal offending patterns (Lewis & Mhlanga, 2001), and probationer behavior (Mackenzie & Li, 2002).

Accuracy, Precision, Validity, and Reliability

Accuracy is commonly defined as the degree to which a measurement represents the true value of the construct under scrutiny. A measurement is said to be precise when results from repeated measurements show a high degree of resemblance. Accuracy is therefore linked to validity, whereas precision is linked to reliability of the measurement instrument. Both LEC and register data are vulnerable to measurement error. Respondents asked to fill in the LEC may experience difficulties in remembering the occurrence or timing of events. Register data on the other hand may be flawed by missing data or mistyped entries. As neither of the available measures is flawless with regard to both prevalence and timing of life-course events, accuracy in the current study thus refers to convergent validity, or the extent to which two *different* measures

of the same construct are in concordance. The word "different" is emphasized here, as some might be inclined to interpret our results in terms of precision, and therefore the reliability of the LEC measure.

Alternate form reliability can be assessed by submitting the same respondents to two equivalent forms of the same measure. In practice, this commonly entails a rewording of the items of the original scale. Precisely when two alternate forms of the same measurement become two different measurements of the same construct may be a matter of debate, yet given the clear differences between the LEC (i.e. retrospective) and the register data (concurrent), we feel confident to interpret our findings in terms of accuracy and validity. The LEC measurement is therefore deemed accurate and valid to the extent that there is a high level of concurrence with the information gathered from the municipal register whereas discrepancies in either the prevalence of events, or the timing of these events would jeopardize its validity.[2]

Sources of Discrepancy

In this section, we elaborate on the sources related to recall errors, which may cause reporting discrepancies. These discrepancies can pertain both to the prevalence as to the timing of specific life events. We draw three sets of hypotheses: (1) hypotheses regarding characteristics related to the specific life events, (2) hypotheses regarding characteristics of respondents in general, and (3) hypotheses regarding personal characteristics more specific to the offender sample. As the main goal of our paper is methodological we keep the theoretical argumentation short.

Characteristics of the events

Recency is a characteristic known to affect reporting discrepancies, as respondents are likely to remember current or recent events better than events in the distant past. The increased ambiguity associated with responses to questions about the past increases the opportunities for selective recall in a manner that enhances the congruity of a response with the respondent's present self-image (Bahr & Houts, 1971). Previous studies regarding memory bias indeed showed that increased length of the retrospective recall period leads to more reporting discrepancies (Auriat, 1993; Sudman & Bradburn, 1973). In line with this research, we pose the following hypothesis:

> H1.1: *Reporting discrepancies are less likely to occur when the event is more recent.*

2. Actually, when the outcomes of two measures known to be subject to error are compared to assess convergent validity as is the case here, discrepancies may signal invalidity of either one of these measures, or both.

Next to recency, the saliency of an event has been noted as an important factor affecting reporting discrepancies (Auriat, 1993; Belli, 1988; Mathiowetz & Duncan, 1988). A salient event is one that is prominent in the respondent's mind and is therefore more likely to be recalled accurately. Saliency is strongly related to the degree to which an event marks a transition point in one's life (Auriat, 1993). Salient events often have economic or social costs and benefits or have continuing consequences after the event (Mathiowetz & Duncan, 1988). Salient events are remembered more accurately because these events signify particular notable transitions in daily routine activities than non-salient events (Belli, 1988). Based on their presumed effect on daily routine activities, we state that one's first marriage, first divorce, and the birth of one's first child, are more salient events than subsequent marriages, divorces, and childbirths. As such we expect these events to be remembered more accurately. More generally, we expect the saliency of events to decline as the number of similar events in the life-course increases. That is, experiencing additional marriages, divorces, and childbirths may negatively influence the accuracy by which these events are reported. Therefore, with regard to saliency, we state the following hypotheses:

> H1.2: *Reporting discrepancies are less likely to occur when the event is the first event of this type experienced.*

> H1.3: *Reporting discrepancies are more likely to occur when the total number of events of the same type experienced over the life-course increases.*

Characteristics of respondents in general

Educational attainment, gender and age are three general personal characteristics that have been found to influence the recall accuracy of self-reported data (Caspi et al., 1996; Morris & Slocum, 2010). The mechanism behind the influence of educational attainment seems straightforward: well-educated might have developed a better memory and are therefore more accurate in remembering life events. We pose the following hypothesis regarding the effect of educational attainment:

> H2.1: *Reporting discrepancies are less likely to occur for respondents with a higher educational attainment, compared to respondents with a lower educational attainment.*

Skowronski and Thompson (1990) found gender differences in recall accuracy. In their study, they found that females date more accurately than males. They ascribe this result to various processes of socialization in the early childhood. The authors, however, are cautious in their conclusion and suggest the need for further research involving differences in the development of temporal schemata, to further examine gender differences in dating accuracy. Psycho-

logical literature indicates that women have a better memory than men, but the exact reasons for this difference are unknown (for a review see: Kimura & Clarke, 2002). Following Skowronski and Thompson (1990), we hypothesize:

> H2.2: *Reporting discrepancies are less likely to occur for female respondents, compared to male respondents.*

Lastly, age is considered an important predictor in recall accuracy, as one's memory is affected by aging (Coughlin, 1990; Levine, Svoboda, Hay, Winocur, & Moscovitch, 2002).

> H2.3: *Reporting discrepancies are more likely to occur when the age of respondents increases.*

Characteristics specific to the offender sample

Frequently changing life circumstances and cognitive difficulties are characteristics specific to offender populations (Schubert et al., 2004). Offenders often live unpredictable lives and month-to-month changes are common in the lives of offenders (Sutton, 2010). Although the LEC method is specifically developed to help respondents with chaotic lives structure their memory and reduce recall errors, a criminal past may signal a lifestyle that increases the risk of recall errors (Engel et al., 2001). The availability of a control group in de CCLS allows us to examine the difference between respondents with an extensive criminal record, respondents with a less extensive criminal record and those without a criminal record in this regard. Therefore, we hypothesize the following:

> H3.1: *Reporting discrepancies are more likely to occur for respondents with a(n extensive) criminal record, compared to respondents without a criminal record.*

Alcohol and drug addiction are common among offenders and chronic substance use can affect cognitive functioning, decreasing the ability of a respondent to comprehend and respond accurately to questions. Moreover, substance use can also impair the ability of the respondent to accurately remember past events (Morris & Slocum, 2010). We therefore test the hypothesis that:

> H3.2: *Reporting discrepancies are more likely to occur for respondents with a history of substance addiction, compared to respondents without a history of substance addiction.*

Data and Measurements

Description of the Dataset

The aim of this paper is to assess the accuracy of data obtained through LECs in an offender sample. The CCLS is well suited to this aim as it consists of a combination

of data obtained through official records and self-reported data obtained through the LEC method. The CCLS is based on a representative 4% sample of all cases of criminal offenses tried in the Netherlands in 1977. The number of cases for drunk driving being very high, the sample for this type of offense was confined to 2%, whereas less common (mainly serious) offenses were oversampled. This sample of cases resulted in an offender sample of $N = 4{,}615$, of which 10% was female. The data also encompass an age-matched comparison group of men not registered for prosecution in 1977 ($N = 741$).[3] Abstracts from the General Documentation Files (GDF) of the Criminal Record Office ("*rap sheets*") were used to reconstruct the entire registered criminal careers of these individuals, covering their lives from age 12 onwards. Criminal records were supplemented with official data on marital and fertility histories, collected from population registers (here after referred to as "*official data*"). At the time of their index offense, the ages of the individuals ranged from 12 to 65, with a mean age of 29 (see Blokland, 2005; Blokland, Nagin, & Nieuwbeerta, 2005 for a detailed description of the dataset).

All (ex-)offenders and comparison persons who were still alive and residing in the Netherlands at the end of 2012 ($N_{total=}3{,}765$: $N_{prosecuted} = 3{,}163$ and $N_{comparison} = 602$) were individually approached between January 2013 and February 2014 to collect self-reported data on several domains of their lives. In total, 959 out of 3,765 respondents participated, resulting in an overall response rate of 25.5%. The response rate among the offender sample was 23.0% ($N = 734$).[4] Among the comparison group, the response rate was 37.7% ($N = 225$).

As the current study is about the dating accuracy of marriage, divorce, and childbirth, only those survey respondents that actually experienced at least one of these events according to the official data were included for further analyses. The sample used consists of 840 survey respondents who were

3. For a random subsample of men from the original CCLS-sample, we chose the subsequent person listed in the military draft registry. If the initial control male happened to be the twin brother of the original CCLS-sample male or if the control male happened to have a criminal case adjudicated in 1977, we chose the next male from the military draft registry. As in the birth cohorts represented in the CCLS-sample all 18-year old men were drafted for military service, the control men can be taken to be representative of the male part of these birth cohorts, who had Dutch nationality, who survived at least to age 18, and did not have a criminal case adjudicated in 1977.

4. A response analysis showed that the offenders that completed the LEC interview did not differ from the offenders that did not participate in the LEC interview in terms of sex, ethnicity, or their average total number of officially registered convictions. We did find that offenders in the LEC sample were on average 5 years older compared to the non-participating offenders. In a prior study, four distinct criminal career trajectories were distinguished in the current sample (Author identifying reference). Using the trajectory group from this study, we find that the proportion of sporadic offenders and medium-rate desisters does not differ between the two subsamples, whereas slight less low-rate desisters and slightly more high rate persisters completed the LEC interview. For the comparison group not convicted in 1977, we find that those who completed the LEC interview did not differ from those who did not in terms of sex, ethnicity, or the likelihood of having been convicted at least once during their life span. Those in the comparison group that did participate in the LEC interview were however about 2 years older on average, and had fewer convictions during the life span than those who did not.

married at least once (N = 766), who were divorced at least once (N = 390), or who had at least one child (N = 719) according to the population register.

The LEC

Computer-assisted personal interviews were conducted using an electronic LEC, which was modeled after the LEC used in the third wave (2008—2009) of the Survey of Health, Aging and Retirement in Europe among respondents aged 50 and older (Schröder, 2011). The respondent's life-course is represented graphically by a grid (or calendar) that is filled automatically during the course of the interview. The grid contained seven rows, representing separate life domains, and columns for each year of the respondent's life from age 12 to their present age. The columns indicated both the calendar year and the respondent's age in that year. In contrast to the conventional interview situation, the interviewer and respondent were sitting next to each other, both facing the computer screen, during the administration of the LEC.

Electronic LECs have several advantages compared to the more widely used paper and pencil LECs. The most important one is that they are less error-prone, as the data feed directly into a data file and do not have to be read and transferred from a paper LEC before analyses. Paper and pencil LECs can become quite messy, and a different person than the interviewer often does data entry. Furthermore, electronic LECs are visually clearer and interactive functions can be built in, which makes them more helpful during the interview (Bellie, 2000).

The interview started with asking names and dates of birth (and death if applicable) of biological and adopted children, followed by the partner history. Start and end years of each relationship were asked, as well as start year of cohabitation, year of marriage and year of divorce. Life *events*, such as marriage, divorce, and childbirth, appeared in the calendar as fully filled cells. The cells between the start and end year of a certain period, for example, the years between the start of a marriage and the year of divorce, were connected by a solid colored line to represent an ongoing situation. Different colors were used for different life domains. By clicking on a particular event in the calendar, additional information appeared below the calendar, such as "birth of {name child}", or "start of marriage to {name spouse}". By clicking on the top cell of a column, all known information about the respondent in that year appeared below the calendar, for example "cohabiting with {name spouse}; moved to Amsterdam; working as carpenter" etc. See Figure 1 for an example of a filled LEC.

As soon as a question about the occurrence of a particular event was answered that event appeared in the LEC and the interviewer could refer to it for help. When, for example, a respondent was not sure about the date of a divorce, this could easily be visually related to the age of the children and changed when necessary. Via such cross-domain stimulation of memory the

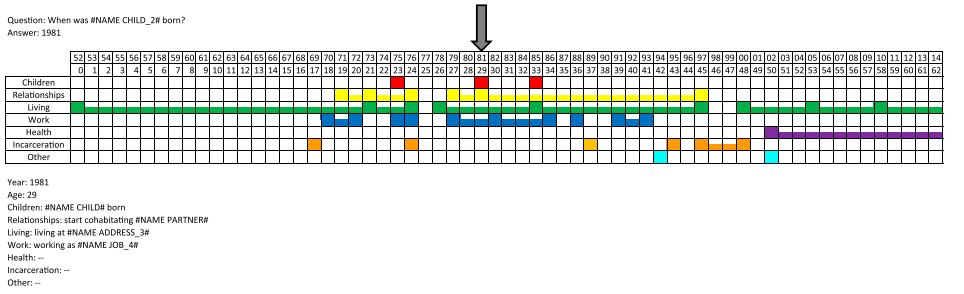

Figure 1 Example of a filled LEC.

LEC is expected to increase dating accuracy (Roberts & Horney, 2009). Moreover, if in answering a question from a subsequent domain it appeared that a previous question was answered incorrectly (e.g. an event was mistimed), this could be corrected. Hence, the interviewer and respondents were working together intensively to fill out the LEC.

Measurements

Dependent variables

There are two sets of dependent variables: (1) dependent variables addressing the reporting discrepancies with regard to the *prevalence* of marriage, divorce, and childbirth, and (2) dependent variables addressing reporting discrepancies with regard to the *timing* of marriage, divorce, and childbirth.

A prevalence reporting discrepancy exists when the number of marriages, divorces, and childbirths differs between the two data sets. For example: when a respondent has four children according to the official dataset, but only mentioned three children in de LEC. To measure reporting discrepancies regarding the prevalence of the life events we thus compared the number of marriages, divorces, and childbirths for each respondent as registered in the official dataset to the number of marriages, divorces, and childbirths a respondent mentioned in the LEC. Three dummy variables were constructed, one for each life event. If the number of a specific life event differs between the two datasets, this respondent was assigned the value 1 on the prevalence variable regarding that life event. If the number of events in the same in both datasets, the respondent was assigned the value 0.[5]

5. Please note that in constructing the prevalence variables, we take the official dataset as benchmark. Marriages, divorces, and childbirth that a respondent mentions in the LEC data, but which do not occur in the official dataset are discarded as these may contain marriages or divorces that have not be formally finalized yet or illegitimate children.

Whereas prevalence reporting discrepancies was measured at the respondent level, timing reporting discrepancies were measured at the event level. A timing reporting discrepancy exists when the year in which a specific event occurred as measured in the LEC is not in concordance with the year a specific event occurred according to the official dataset. In order to construct the dependent variables measuring timing reporting discrepancies, we compared the year of each marriage, each divorce, and the birth of each child as mentioned by respondents in the LEC to the year of each marriage, each divorce, and birth of each child as recorded in the official dataset. The resulting variables are dichotomous, with value 1 indicating that a different year was mentioned in both datasets and value 0 indicating that the same year was measured in both datasets.

Note that when information on a marriage, divorce, or childbirth was missing in the LEC data, this specific life event was excluded from the analysis with regard to *timing*. These recall errors are found in the analyses regarding *prevalence*. For instance, consider a respondent who, according to the official data, has been married two times, in 1960, 1965, and in 1970. According to the LEC, this respondent has been married two times, once in 1965 and once in 1970. This respondent will have a reporting discrepancy with regard to the prevalence of marriage, but none regarding the timing of marriage.

Independent variables

At the event level, we constructed the following recency variables: (1) *years since marriage*, (2) *years since divorce* and (3) *years since childbirth* for each of the marriages, divorces, and childbirths that occurred in a respondent's life. These variables were constructed to count backwards from the year of interview (2013 or 2014) to the year that the specific event occurred according to the official dataset. A higher score on this variable indicates that the event occurred longer ago than a lower score on this variable. In order to measure saliency, a dummy variable *first time event* was constructed indicating the first marriage, divorce, and childbirth. At the respondent level saliency was operationalized by a measure indicating for each event type, the total amount of similar events the respondent experienced during his or her life-span up to the time of the interview *(number of events*).

At the respondent level, *educational attainment* was measured through respondents' highest level of education. The answers were recoded into primary education or less (0), secondary education (1), and tertiary education or more (2). *Gender* was coded 1 if the respondent is male and 0 if the respondent is female. A variable *age* was created measuring the age of respondents at the time of the interview. Criminal records were used to operationalize the variable *criminal record*. We constructed ordinal variable indicating the value 0 for respondents who have never been convicted for a criminal offense, 1 for respondents who have been convicted for a criminal offense less than four

times, and 2 for respondents who have been convicted for a criminal offense four times or more. The variable *past substance addiction* was operationalized through the following questions (translated from Dutch): "Did you experience one or more periods in your life in which you were addicted to drugs?" and "Did you experience one or more periods in your life in which you were addicted to alcohol?" Answer categories were 0 (no I did not), 1 (yes, once) and 2 (yes, more than once). We constructed a dichotomous variable indicating 1 if respondents answered 1 or 2 to either of these two questions and 0 otherwise. A description of the sample can be found in Tables 1 and 2.

Finally, we controlled for whether a respondent was incarcerated during the year that an event occurred. For each life event, a dummy variable was constructed which indicates 1 if the respondent was incarcerated in the year the event occurred according to the official data, and 0 if the respondent was not incarcerated in the year the event occurred according to the official data. This variable was included as we suspected that incarceration may act as a time marker organizing memories and thereby leading to fewer discrepancies.

Table 1 Sample description: respondent characteristics ($N = 840$[1])

	Range	Proportion/mean (std.)	*N*
Educational attainment			
Primary	0—1	.27	229
Secondary	0—1	.53	445
Tertiary	0—1	.20	166
Gender			
Female	0—1	.07	59
Male	0—1	.93	781
Criminal record			
0 Offenses	0—1	.22	178
1—3 Offenses	0—1	.33	276
4 Offenses or more	0—1	.46	386
Substance addiction			
No	0—1	.75	634
Yes	0—1	.25	206
Age	48—90	60.39 (7.23)	840
Number of marriages	0—6	1.19 (.69)	840
Number of divorces	0—6	.56 (.72)	840
Number of children	0—9	1.98 (1.27)	840

[1]This table pertains to the *entire* sample. The subsamples regarding marriage, divorce, and childbirth are smaller as in these subsamples all respondents without this specific life event are excluded. Please note that the overall distribution of the respondent characteristics does not vary significantly between the subsamples.

Table 2 Sample description: event characteristics[1]

	Range	Mean (std.)/proportion	*N*
Years since event			
Marriage	0–63	29.90 (11.98)	927
Divorce	0–47	21.27 (11.24)	413
Children	1–62	28.76 (10.60)	1,592
First time event			
Marriage	0–1	.78	927
Divorce	0–1	.84	413
Children	0–1	.42	1,592

[1]Based on LEC data, only for those events reported in both the official data and the LEC data.

Analytical Strategy

Before turning to the multivariate test of our hypotheses, we will first provide some descriptive statistics regarding the occurrence of prevalence and timing discrepancies in the data. We will proceed by presenting bivariate snapshots of the data which provide an indication as to how these discrepancies are distributed amongst the characteristics of interest. For reasons discussed below, we will continue to test our hypotheses in a multivariate way only for timing reporting discrepancies. As the data for timing discrepancies have a nested structure with events nested within respondents and due to the binary nature of the dependent variable, multilevel logistic analyses are needed to accurately test the hypotheses. Separate analyses were conducted to examine the determinants of timing discrepancies regarding marriage, divorce, and childbirth. We present both the coefficients as well as the odds ratios. As the results of logistic multilevel models can be difficult to interpret, we additionally provide predicted probabilities for the significant effects. The probabilities were obtained by running the postestimation command "predict" in Stata. This command calculates the linear predictor term for our model regarding a specific life event. The "tabstat" command was subsequently used to obtained the predicted probabilities for specific groups, keeping all other variables at their means.

Results

Descriptive Results Regarding Reporting Discrepancies

Table 3 presents the frequency distributions of marriage, divorce, and childbirth among respondents according to the official dataset and according to the self-reported LEC data. We see that there are relatively few respondents who have been married or divorced three times or more, and there are few who

Table 3 Distribution of marriages, divorces and childbirths per respondent according to the official municipal data and according to the self-reported LEC data (*N* = 840)

	Official municipal data		LEC data	
	%	*N*	%	*N*
Marriage				
Not married	8.81	74	10.24	86
1	68.10	572	69.52	584
2	18.57	156	17.14	144
3	3.93	33	2.62	22
4 or more	.60	5	.48	4
Divorce				
Not divorced	53.57	450	55.15	463
1	39.29	330	35.71	300
2	5.12	43	7.74	65
3	1.79	15	1.19	10
4 or more	.24	2	.24	2
Childbirth				
No children	13.45	121	11.43	95
1	17.26	140	16.19	136
2	42.26	353	41.43	348
3	16.79	139	18.69	157
4	6.79	57	7.98	67
5	2.14	17	2.50	21
6 or more	1.31	13	1.79	16

have five children or more. The comparison between the two data sources shows us that according to the LEC data, more respondents are married at least once, fewer respondents have been divorced at least once, and fewer respondents have at least one child, than according to the official data. Based on Table 3, we can say that reporting discrepancies exist regarding the prevalence of all three life events.

Table 4 shows the percentages of prevalence and timing reporting discrepancies regarding the three life events. It is important to note that prevalence reporting discrepancies are based on events which were not reported in the LEC data, but which did occur according to the official dataset.[6] Timing reporting discrepancies are only calculated for those events for which we have data

6. Discarding events that were reported on the LEC, but not in the official data, we de facto use the official data as a benchmark against which to compare the validity of the LEC data. In total 28 respondents who were continuously single according to official data reported to have married at least once, 26 reported to have experienced divorce while there was no divorce registered in the official data, and 78 respondents reported to have parented at least one child, whereas no children were registered in the official data.

Table 4 Proportion of prevalence and timing discrepancies with regard to marriage, divorce and childbirth

	Prevalence reporting discrepancy[1] (respondent level)			Timing reporting discrepancy (event level)		
	N_{total}	%	(*N*)	N_{total}	%	(*N*)
Marriage	766	8.88	(68)	927	27.51	(255)
Divorce	390	14.62	(57)	413	67.31	(278)
Childbirth	719	5.72	(41)	1,592	10.93	(174)

[1]Note that in the prevalence reporting discrepancies, respondents who never experienced an event according to both data sources (e.g. respondents who correctly report that they have never married), are left out. If they would have been included, the percentages with a discrepancy would be lower.

from both the official and the LEC dataset. Table 4 shows that both prevalence and timing reporting discrepancies exist. However, prevalence reporting discrepancies occur less than timing reporting discrepancies. Overall, reporting discrepancies regarding divorce occur most often (prevalence: 14.62% and timing: 67.31%). Of the 390 divorced respondents in the official dataset, 57 respondents (14.62%) mentioned a different number of divorces in the LEC than was given in the official dataset. Furthermore, of the 413 divorces that were mentioned in both datasets, 278 (67.31%) mistimed the specific year of that divorce.

Figure 2 provides the distribution of timing discrepancies with regard to marriage, divorce, and childbirth. Please note that for this figure only, timing reporting discrepancies are measured by the size of the discrepancy in years rather than as dichotomous variable. Cases of backward telescoping—reporting events as more remote than registered in the official data—are depicted as positive timing discrepancies, whereas cases of forward telescoping—reporting events as more recent—are depicted as negative. As Figure 2 shows, in about 40–50% of the events for which there is a timing reporting discrepancy, respondents report the particular event taking place within one year—either before or after—the year it was registered in the municipal data. The data regarding marriage show that 150 (58%) of the timing reporting discrepancies deviate by 1 year and 50 (19%) of the reporting discrepancies deviate by 2 years. Regarding divorce, 109 (39%) of the timing reporting discrepancies deviates by one year and 53 (19%) deviate by two years. Of the timing reporting discrepancies regarding childbirth, 125 (59%) deviate by one year and 39 (19%) deviate by two years.

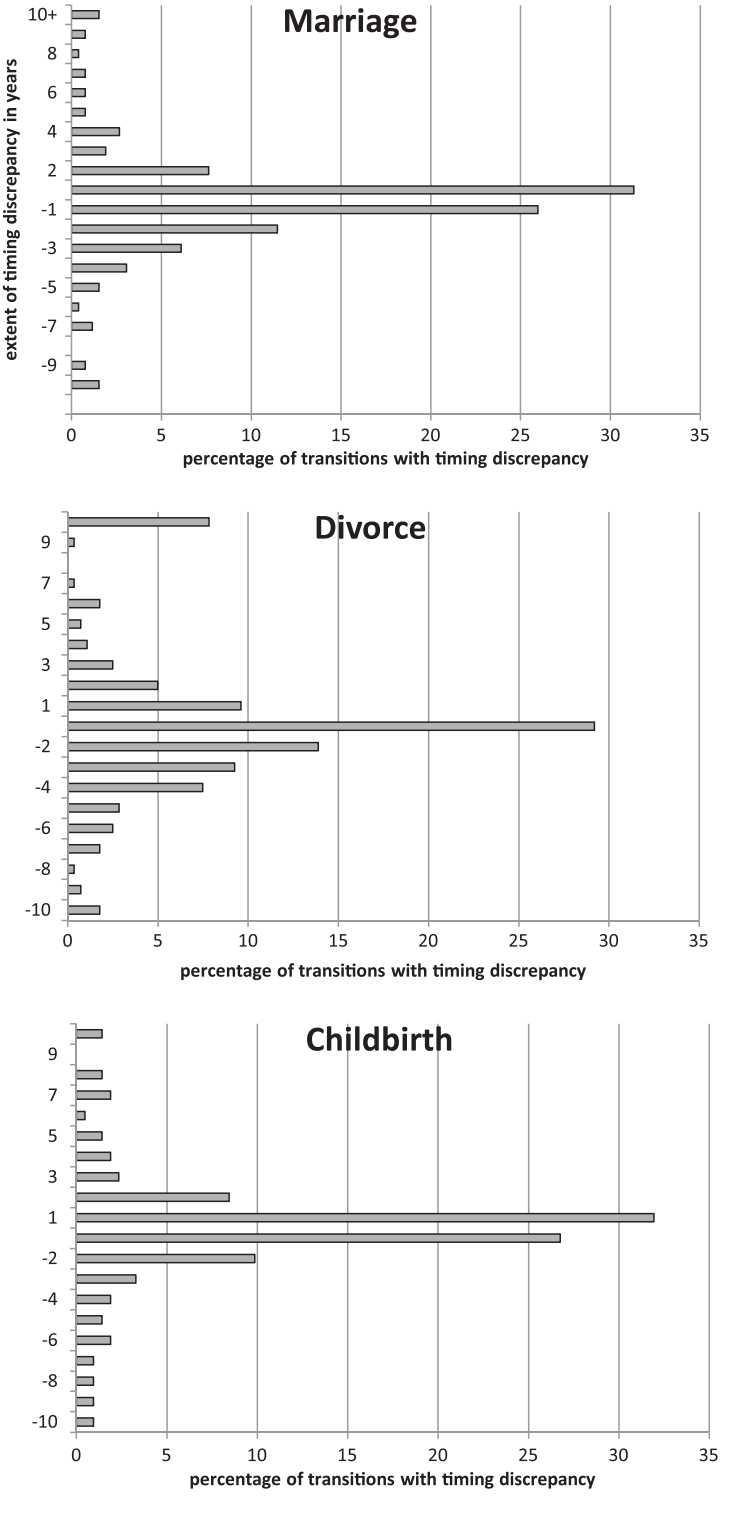

Figure 2 Distribution of timing reporting discrepancies regarding marriage (*N* = 262), divorce (*N* = 281), and childbirth (*N* = 213), by size of the discrepancy in years.

Bivariate Analyses Regarding Prevalence Reporting Discrepancies

As concluded in Table 4 prevalence reporting discrepancies exist to a lesser extent than timing reporting discrepancies. More specifically, the dataset contains too few prevalence reporting discrepancies to allow for a multivariate test. For prevalence reporting discrepancies we will therefore only present bivariate results. In this section, we discuss the distribution of prevalence reporting discrepancies by educational attainment, gender, age, and substance addiction. Figure 3 shows the percentages of prevalence reporting discrepancies for each life event by characteristics of the respondents. We indicate whether differences between groups are significant. Because of the small sample sizes (see Table 4) we also report which differences are significant at an alpha-level of .10.

In Figure 3(a) we see that, for all three life events, prevalence reporting discrepancies occur most often for respondents whose highest level of education is primary education or less. The difference between these respondents and the respondents whose highest level of education is secondary or tertiary is statistically significant for all three life events. There are no significant differences in prevalence discrepancies between respondents with secondary and tertiary education for all life events. Hence, we found partial support for Hypothesis 2.1 predicting that prevalence reporting discrepancies are less likely to occur for respondents with a higher educational attainment, compared to respondents with a lower educational attainment.

The difference in percentages reporting discrepancies between males and females is depicted in Figure 3(b). Males seem to underreport their number of divorces and children more often than females. Females, on the other hand, seem to underreport their number of marriages more than males. However, these differences are not significant. Hence, although the patterns seem to be in line with Hypothesis 2.2 for divorces and childbirths, we cannot confirm Hypothesis 2.2.

Figure 3(c) shows the percentage of reporting discrepancies by the age of respondents. For all three life events, reporting discrepancies seem to be more likely to occur for the oldest respondents compared to the youngest respondents. However, significant differences are only found regarding childbirth. Regarding marriage and divorce no significant differences exist between any of the age groups. Regarding childbirth, the oldest age group is more likely to have a prevalence discrepancy than each other of other age groups, but no differences are found amongst the other age groups. Hence, we can only partially confirm Hypothesis 2.3, stating that age negatively influences reporting discrepancies.

Figure 4 shows the prevalence reporting discrepancies regarding the life events by the characteristics specific to the offender population: criminal record and past substance addiction. As can been seen in Figure 4(a), the data are in line with Hypothesis 3.1: reporting discrepancies occur significantly more amongst respondents who have an extensive criminal record (4 offenses

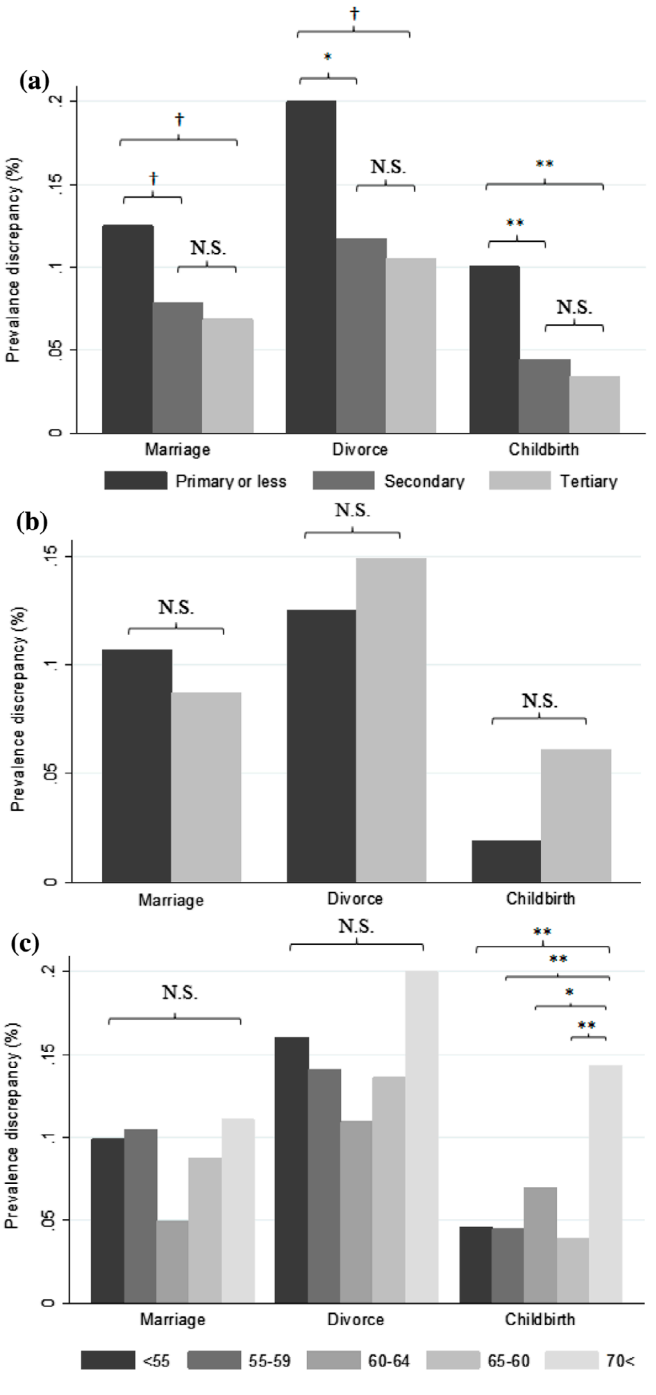

Figure 3 Prevalence reporting discrepancies regarding marriage, divorce, and childbirth by characteristics of the respondent.
Notes. $^{\dagger}p < .1$, $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

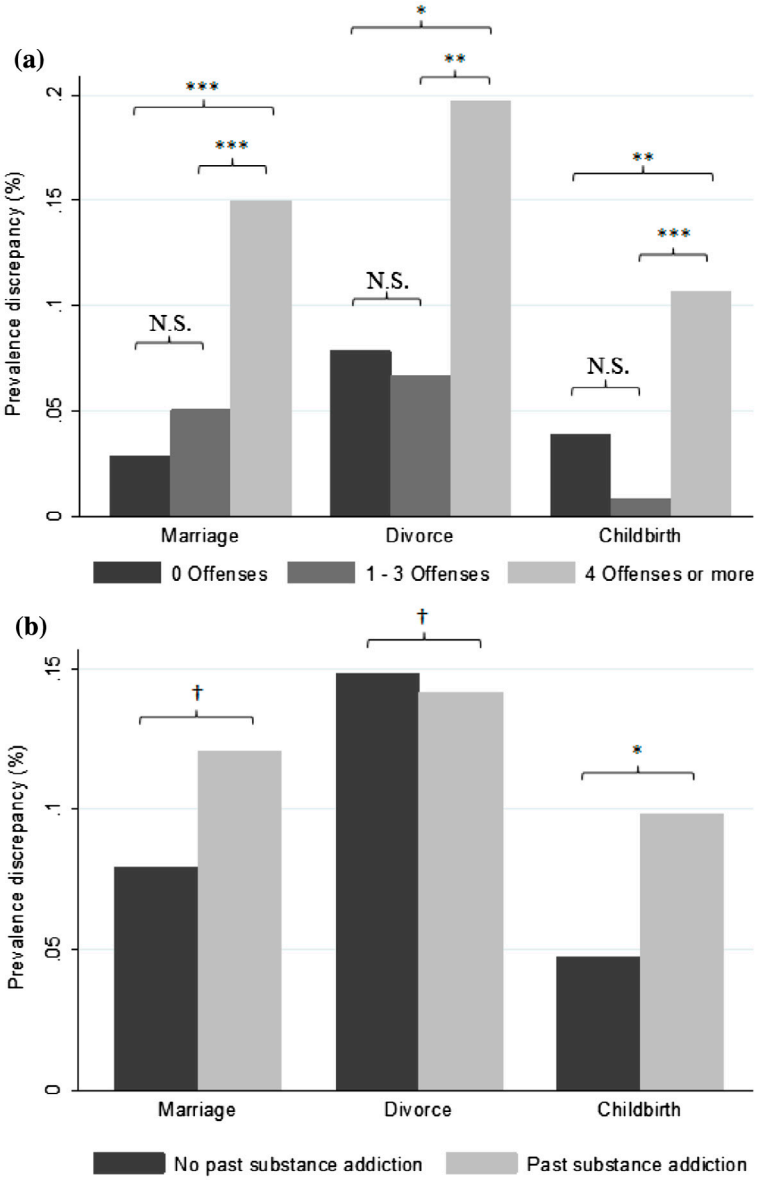

Figure 4 Prevalence reporting discrepancies regarding marriage, divorce, and childbirth by characteristics of the respondent specific to the offender sample.
Notes. $^{\dagger}p < .1$, $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

or more) than amongst respondents who do not have a criminal record or a more limited criminal record (1—3 offenses).

Hypothesis 3.2 predicts that prevalence reporting discrepancies are more likely to occur for respondents with past alcohol or drug addiction than for

respondents without past alcohol or drug addiction. Figure 4(b) shows that for all three life event the data seem to be in line with this hypothesis.

Bivariate Analyses Regarding Timing Reporting Discrepancies

Figure 5 presents the bivariate patterns regarding timing reporting discrepancies based on respondent characteristics. Figure 5(a) shows that timing reporting discrepancies regarding marriage and childbirth are less likely to occur when the educational level of the respondent increases. This pattern seems to be the same for divorce, but in this case the differences between the educational groups are not significant. Figure 5(b) relates to gender and illustrates that males seem to be more likely to mistime the birth of their children than females. Although males also seem to be slightly more likely than females to mistime their marriages and divorces, differences with respect to mistiming these events are not significant. In Figure 5(c), we see that although there are some small differences between specific age groups, age does not seem to affect the mistiming of marriage, divorce, or childbirth to a large extent. Please note that in Figure 5(c) only the significant differences are marked, we did not add markers for each of the non-significant age group differences as that would result in a messy picture.

Figure 6 presents the bivariate patterns regarding timing reporting discrepancies based on respondent characteristics specific to the offender sample. Both respondents who have been convicted for a criminal offense and respondents with a past substance addictions are significantly more likely to mistime on all three life events.

It is important to underline that Figures 5 and 6 only present bivariate differences between two groups without controlling for any of the other independent variables. The overall higher number of timing discrepancies, as compared to prevalence discrepancies, allows for a more elaborate test of our hypotheses applying multivariate models. Results from these models are presented in Table 5.

Multilevel Analyses Explaining Timing Reporting Discrepancies

Marriage

Model 1 of Table 5 presents the multilevel model estimated to test the hypotheses regarding the occurrence of timing reporting discrepancies for marriage. The underlying multilevel model has two levels: the event level ($N = 927$) and the respondent level ($N = 743$). The Intra Class Correlation (ICC) of the intercept only model (not presented: $\rho = .16$) indicates that 16% of the total variance is represented at the respondent level.

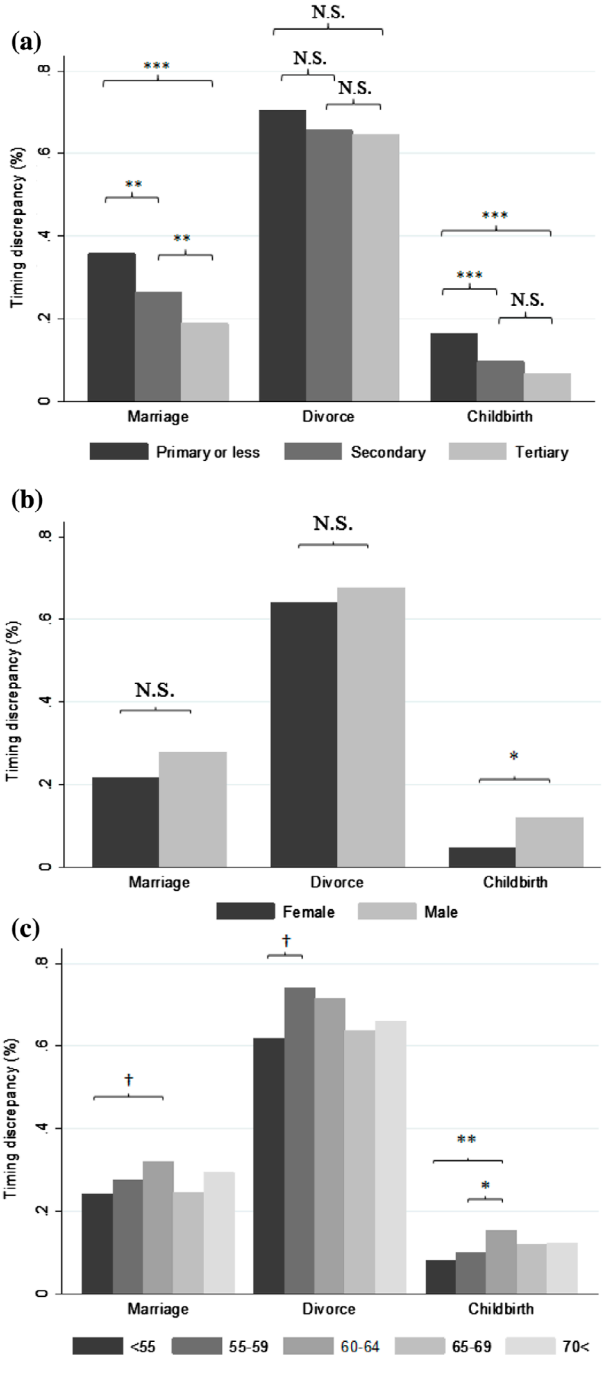

Figure 5 Timing reporting discrepancies regarding marriage, divorce, and childbirth by characteristics of the respondent.
Notes. $^{\dagger}p < .1$, $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

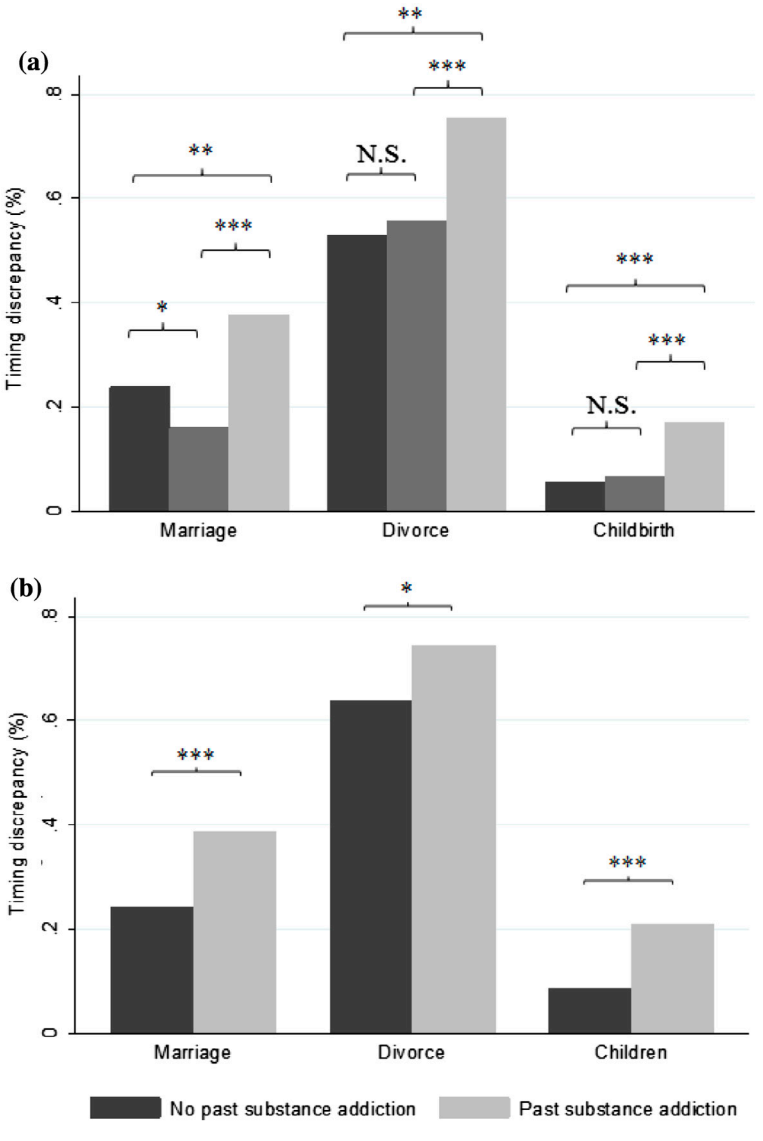

Figure 6 Timing reporting discrepancies regarding marriage, divorce, and childbirth by characteristics of the respondent specific to the offender sample.
Notes. $^{\dagger}p < .1$, $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

Regarding the characteristics of the event, results from Model 1 show that only the total number of marriages has a significant influence on the likelihood that a timing reporting discrepancy exits regarding marriage. Timing reporting discrepancies regarding marriage are not significantly less likely to occur for more recent marriages ($b = -.00$, $p = .67$) or for the first marriage ($b = .22$, $p = .43$). Therefore, Hypothesis 1.1 and Hypothesis 1.2 are not confirmed regarding marriage. Based on Model 1, we can confirm Hypothesis 1.3: timing

Table 5 Results logistic multilevel models with dependent variable timing reporting discrepancies

	Marriage (*N* = 927 events nested in 743 individuals)			Divorce (*N* = 413 events nested in 359 individuals)			Childbirth (*N* =1,592 events nested in 715 individuals)		
	Model 1			Model 2			Model 3		
	Coef.	S.e.	OR	Coef.	S.e.	OR	Coef.	S.e.	OR
Fixed part									
Constant	−3.41***	1.01	.03	.39	1.40	1.48	−6.61***	1.56	.00
Independent variables									
# Years since event	−.00	.01	1.00	.05**	.02	1.05	.03	.02	1.03
First time event	.22	.28	1.25	−.42	.42	.65	.60*	.23	1.82
Total # of events	.33*	.13	1.39	−.32	.32	.73	.42***	.11	1.52
Education (ref=primary or less)									
Secondary education	−.06	.20	.94	−.04	.29	.96	−.16	.32	.85
Tertiary education	−.42	.42	.65	.15	.41	1.16	−.15	.44	.86
Male	.27	.36	1.35	.24	.40	1.27	1.12	.66	3.05
Age	.03	.01	1.03	−.01	.02	.99	−.01	.03	.99
Criminal record (ref = 0)									
1—3 Offenses	−.62*	.25	.54	.13	.40	1.14	−.07	.43	.93
4 Offenses or more	.23	.24	1.33	1.11*	.46	3.03	.69	.43	2.00
Substance addiction	.47*	.20	1.60	.26	.29	1.30	1.23***	.33	3.43
Control variables									
Incarcerated year event	.28	.33	1.33	−.94*	.46	.39	1.05**	.40	2.86
Random part									
Variance (event)	3.29			3.29			3.29		
Variance (respondent)	.24	.40		.51	.93		2.99	.90	
Fit statistics	Parameter	Df		Parameter	Df		Parameter	Df	
Deviance	1,022.83	13		477.97	13		931.67	13	

*$p < .05$; **$p < .01$; ***$p < .001$.

reporting discrepancies are more likely to occur when the total number of marriages of respondents increases. The predicted chances of mistiming increases from approximately 22% for respondents who have been married once to approximately 39% for respondents who have been married three times or more. Model 1 shows that timing reporting discrepancies regarding marriage do not vary with highest level of completed education (secondary: $b = -.06$, $p = .74$, tertiary: $b = -.42$, $p = .12$). Therefore, we cannot confirm Hypothesis 2.1. From Model 1, we cannot confirm Hypothesis 2.2 or Hypothesis 2.3, which predict that timing reporting discrepancies are less likely to occur among female and younger respondents than among male ($b = .30$, $p = .36$) and older respondents ($b = .03$, $p = .09$). Interestingly, respondents with a criminal record do not mistime their marriage significantly more often than respondents without a criminal record. In contrast to Hypothesis 3.1, respondents who were convicted for less than four offenses have a predicted chance of approximately 15% to mistime their marriage, in contrast to a chance of approximately 23% for respondents who never have been convicted for an offense. In line with Hypothesis 3.2, results show that respondents who have a history of substance addiction are more likely to mistime their marriage, compared to respondents who do not have a history of substance addiction. The predicted chance of mistiming increase from approximately 23% for respondents who never experienced a period of substance addiction to approximately 38% for respondents who did. The model controls for the effects of incarceration at the time of marriage, but this does not affect the misting of marriage.

Divorce

Model 2 of Table 5 presents the multilevel model estimated to test the hypotheses regarding the occurrence of timing reporting discrepancies for divorce. The sample exists of 413 divorces nested in 359 respondents. The ICC of the intercept-only model (not presented: $\rho = .26$) indicates that roughly 26% of the total variance is represented at the respondent level.

In line with Hypothesis 1.1, Model 2 shows that recency has a significant positive effect on mistiming. Divorces that occurred longer ago are recalled less accurately. The predicted chance of mistiming a divorce increases from approximately 42% to approximately 67% when the number of years since the divorce took place increases from one to 15. In line with our findings for marriage (Model 1), the first divorce is not mistimed less often than subsequent divorces ($b = -.42$, $p = .45$). Also the number of divorces does not significantly affect the existence of a timing reporting discrepancy ($b = -.32$, $p = .22$). Therefore, we cannot confirm Hypothesis 1.2 and Hypothesis 1.3. None of the general respondent characteristics significantly influences the likelihood of a timing reporting discrepancy regarding divorce. Respondents with a lower level of education, males, or older respondents do not mistime their divorce more often than respondents with a higher level of education (secondary education:

$b = -.04$, $p = .90$, tertiary education: $b = .15$, $p = .71$), females ($b = .24$, $p = .54$), or younger respondents ($b = -.01$, $p = .55$). Therefore, we cannot confirm Hypothesis 2.1, Hypothesis 2.2, and Hypothesis 2.3.

Respondents who have been convicted for a criminal offense four times or more are significantly more likely to mistime their divorce than respondents who have never been convicted for a criminal offense. The predicted chance of mistiming a divorce increases from approximately 53% for respondents who never have been convicted for a criminal offense to approximately 77% when respondents have been convicted for more than three offenses. This confirms Hypothesis 3.1. In contrast to Model 1 for marriage, substance addiction does not have an effect on reporting discrepancies regarding divorce ($b = .26$, $p = .37$). Based on Model 2, we cannot confirm Hypothesis 3.2. Model 2 controls for the effects of incarceration at the time of divorce. Respondents incarcerated in the year their divorce took place mistime their divorce less often compared to respondents not incarcerated in the year their divorce took place.

Childbirth

Model 3 presents the multilevel model estimated to test the hypotheses regarding the occurrence of timing reporting discrepancies for childbirth. The sample contains the birth 1,592 children nested in 715 respondents. The ICC of the intercept only model (not presented: $\rho = .70$) indicates that roughly 70% of the total variance is represented at the respondent level.

Model 3 shows that recency does not significantly influence the mistiming of childbirth ($b = .03$, $p = .08$). Therefore we cannot confirm Hypothesis 1.1 regarding childbirth. In contrast to Hypothesis 1.2, Model 3 shows that the birth of a first child is more often mistimed than the birth of subsequent children. In line with our model regarding marriage, saliency as defined by the total number of births positively influences the mistiming of childbirth. The predicted chance of mistiming increases from approximately 6% for respondents with one child to approximately 11% for respondents with four children or more. This confirms Hypothesis 1.3 regarding childbirth. In line with the model regarding marriage and divorce, there are no significant effects of level of education (secondary education: $b = -.16$, $p = .62$; tertiary education: $b = -.15$, $p = .73$), gender ($b = 1.12$, $p = .09$), or age ($b = -.01$, $p = .72$) on the likelihood of a timing reporting discrepancy regarding childbirth. Therefore, we are not able to confirm Hypothesis 2.1, Hypothesis 2.2, and Hypothesis 2.3. In contrast to the results regarding divorce, respondents with a criminal record did not mistime the birth of their children significantly more than respondents without a criminal record. Therefore, Hypothesis 3.1 cannot be confirmed regarding childbirth. Past substance addiction is positively related to the mistiming of childbirth. In line with Hypothesis 3.2, respondents with a past substance addiction have a chance of approximately 17% to mistime the birth of one of their children, compared to a chance of approximately 6% for respon-

dents without a past substance addiction. Model 3 controls for the effects of incarceration at the time of childbirth. Respondents incarcerated in the year one of their children was born mistime this birth more frequently compared to respondents not incarcerated in the year of childbirth.

Clustering of timing reporting discrepancies within respondents

Up to this point, we took the event as unit of analysis and examined whether characteristics of the event itself or of the respondents experiencing these events influenced the occurrence of reporting discrepancies. However, since multiple events are often embedded in the same respondent, a small group of respondents, mistiming many or even all events they experienced, could be responsible for a large part of the total number of timing reporting discrepancies found for the entire sample. If many discrepancies originate from a limited number of respondents, eliminating these respondents from the sample would result in a sample with more accurate LEC data. In order to examine the extent to which timing discrepancies are concentrated in a small subgroup of respondents, we constructed a variable which counts the number of reporting discrepancies made by each respondent. Figure 7 shows the clustering of reporting discrepancies within respondents.

Figure 7 plots the percentage of the sample against the percentage of discrepancies and resembles the well-known Lorenz curve. In criminology, Lorenz curves are commonly used to depict the concentration of crimes across for instance individuals, families, or geographic areas (Bernasco & Steenbeek, 2017; van de Weijer, Thornberry, Bijleveld, & Blokland, 2015). Here, we similarly examine the concentration of reporting discrepancies across individuals. If all respondents were to contribute equally to the total number of reporting discrepancies, the plot would show a straight diagonal line from the bottom left to the top right of the graph (the reference line). The more the distribution of reporting discrepancies is skewed, the more the observed line deviates from the diagonal reference line. As becomes clear from Figure 7 discrepancies do cluster within respondents—respondents reporting one discrepancy tend to report others—but only to a limited extent. The distribution of discrepancies regarding childbirth is most skewed with 25% of the respondents being responsible for 45% of all reported discrepancies regarding this event. Skewness is less pronounced for marriage and divorce. This in part reflects the more limited range for the total number of marriages and divorces per respondent compared to the maximum number of children, which in turn limits the maximum possible number of reporting discrepancies per respondent. Finally, the finding that the curve summing discrepancies across all three types of events is most skewed indicates that respondents, who, for instance, mistime the year of their marriage, are more likely to also mistime the year of their divorce and/or the year they became a parent, compared to respondents that timed their marriage(s) correctly.

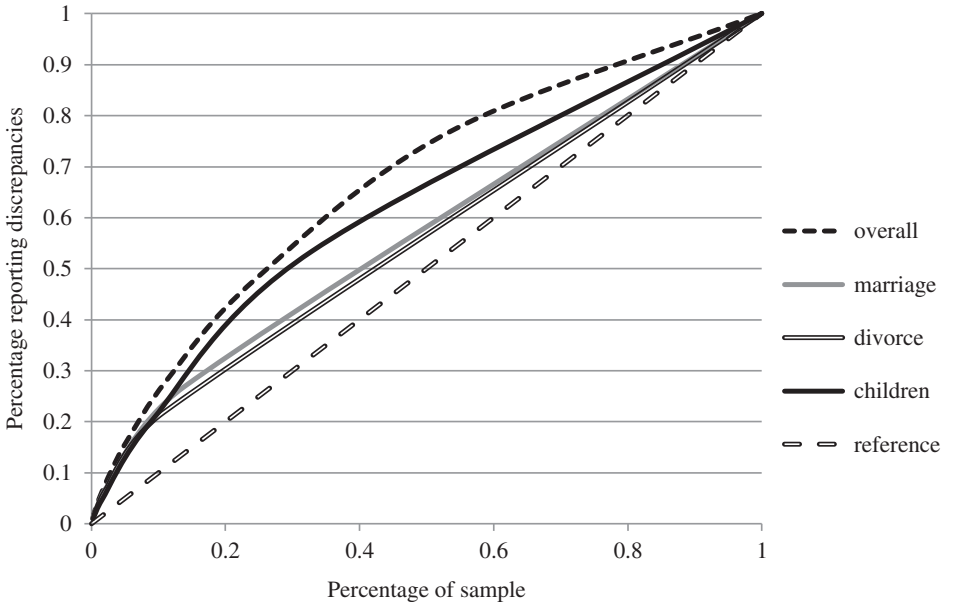

Figure 7 Clustering of timing reporting discrepancies within respondents for each of the life event.

Discussion and Conclusions

In this study, we aimed to assess the accuracy of self-reported data on key life events obtained through the LEC method in an offender sample. LEC methods are increasingly used as a way of collecting retrospective data that are more valid than ordinary self-reported survey data (Sutton, 2010). Most importantly, because the LEC method claims to reduce recall errors as the interviewer and the respondent work together intensively to fill out, and correct, the LEC (Roberts & Horney, 2009). However, research examining how accurate these data truly are is scarce, especially among criminologically relevant samples. In the present study we compared offenders' self-reported data on marriage, divorce, and childbirth to official municipality data and examined whether and to what extent reporting discrepancies occurred.

When we compare the *number* of marriages, divorces, and children respondents have according to the official municipality dataset to the number of marriage, divorces, and children respondents themselves indicated to have in the LEC, we can conclude that prevalence reporting discrepancies are sparse; the LEC method thus is able to obtain quite valid data on these aspects. However, the purpose of using an LEC is especially to acquire reliable data on the *timing* of life-course events, in order to enable life course research to make causal inferences with regard to the effects these time varying variables have on criminal development and vice versa. When we compare the actual year of marriage, divorce, and childbirth as indicated in the LEC to the year as given

in the official municipality dataset, results clearly show more reporting discrepancies. In about 28% of the marriages, 68% of the divorces and 11% of the childbirths reported in both datasets, there was a discrepancy in the year of the event.

When interpreting these percentages it is important to consider that the average age of our respondents was 60 at the time of the LEC-interview and that we asked them to reflect on their entire life span. Prior studies using the LEC-method usually span shorter periods, though life-time LEC have been used before (e.g. Porcellato, Carmichael, & Hulme, 2016). We conclude that timing discrepancies are least common with regard to childbirth, and most common with regard to divorce. This may reflect the fact that the birth of a child is a clearly demarcated event, which is often commemorated annually from the event onwards increasing the likelihood of correct recollection. Divorce, and to a lesser extent marriage, is less clearly demarcated as partners can cohabit without being married, separate without being divorced, and even divorce without separation. Respondents might therefore mistakenly report the year of separation instead of the year the divorce was officially registered by the municipality. The finding that almost half of the reporting discrepancies fell within a one year boundary from the registered event seems to fit this interpretation.

When looking at the timing reporting discrepancies more closely, we find that characteristics specific to an offender sample—an extensive conviction history, and self-reported alcohol or drug addiction—most often influence the likelihood of mistiming events. A conviction history is associated with more mistiming on divorce, whereas substance addiction affects timing discrepancies regarding both marriage and childbirth, but not divorce. While this may reflect the notion that offenders' lives are often "chaotic and unstable", this chaos and instability would have to go beyond that what is already captured by the other variables in the model.

Recency influenced the accuracy of timing of divorces—reporting discrepancies regarding divorce were less likely when they took place more recent, but not marriage or childbirth. When interpreting this finding, one has to take into account that, due to the age of our respondents, relatively few marriages and childbirths occurred within the ten years preceding the interview. If our sample would have contained more recent marriages and childbirths, in addition to marriages and childbirths that occurred longer ago, perhaps recency effects would have been found for this event as well. Findings for saliency are mixed. Whether the event was a first marriage or divorce did not affect the accuracy of timing of events. Interestingly, the birth of the first child was more often mistimed than the birth of subsequent children. The total number of events experienced did influence mistiming, however, be it only for marriage and childbirth. Respondents experiencing more marriages and more childbirths over their life span, were also more likely to mistime these events.

Educational attainment did not influence timing reporting discrepancies. Male and female respondents did not differ in their likelihood to mistime

events. The lack of significance in this regard could in part be due to the limited number female offenders reporting these life events in our sample. A more substantive interpretation could be that compared to male offenders, female offenders are characterized by higher levels of mental health problems (Hollin & Palmer, 2006). This may nullify the advantage woman in general have over males when it comes to remembering key life-course events in this particular population.

How do our results on the self-reported prevalence and timing of key life events speak on the use of LECs as a data collection tool to gather life-course information in an offender population? Whereas reports on the prevalence of events are mostly accurate, reports on the timing of events are so to a lesser extent. Given that inaccuracies in the timing of events often do not exceed one year prior or after the registered event, the observed level of timing discrepancies for both marriage and childbirth seems unlikely to substantively affect prior research using the LEC to an extent that it threatens the conclusions reached in these studies. Many of these studies have focused on the effects of "being married" or "being a parent" in a particular year, and not so much on the event of marriage or childbirth per se. Even if respondents mistime the exact year of marriage or childbirth, their report of being married or being a parent will be accurate in the majority of the years reported upon. More recently however, researchers have begun to scrutinize the development of crime in the years surrounding these key life events, hypothesizing that the decline in offending often observed following such an event, is in fact already present in the years or months preceding the event (Skardhamar & Savolainen, 2014). Based on the current results, it seems unlikely that data gathered with an LEC that covered the entire life span in a sample of adult offenders has the level of accuracy to sustain this type of analysis.

What do these results on marriage, divorce, and childbearing suggest about the validity of LEC data, collected among offenders, on other life-course domains such as housing, health or employment? A general assumption behind the design of LECs is that events such as marriage and childbirth are generally quite important and salient events in people's lives, and therefore easier to remember and time accurately than other events, such as moving house or starting a new job. In LEC interviews childbirths and marriages are asked first so that they can serve as a capstone to remember the timing of other events. On the one hand therefore, one could expect that retrospective data on other life-course domains, like employment, even when collected with LECs will not be more or even as accurate than those collected for children, marriage, and divorce, based on the saliency of these latter events in the life-course. On the other hand, our results may also reflect the "fuzziness" or certain transitions: what is experienced as a "divorce", may not always coincide with what is officially registered as such. In case of, for instance, employment, transitions may be more easily demarcated—if one is fired or resigns, this usually coincides with actually not going to work anymore—which may lead to transitions in this domain being more accurately reported. Whereas respondents in the CCLS

were questioned on their employment histories using the LEC, we do not have access to official data on respondents' employment histories. Future research, using a different sample, will therefore be needed to yield a decisive answer on this point.

One possible way to increase the validity of data gathered through the use of an LEC might be to provide respondents with "anchors" taken from available register data, either prior to, or after they have completed the LEC on a particular topic. When confronted with register data on, for instance, his official divorce date, a respondent may either come to recognize his timing error, and revise his LEC data accordingly, or might provide the interviewer with valuable information on the reasons behind the apparent discrepancy. Even if official data is not available for the entire period under scrutiny, providing official anchors during the period that these data are available, may help respondents in accurately filling in the LEC for the remaining period. Another option would be to try to identify respondents that have difficulty remembering important events and transitions in their lives, for example based on the impression respondents make on the interviewer, and discard these respondents from further analysis. However, besides that it might be difficult for interviewers to judge the accuracy of the information provided by the respondent, we find that the distribution of reporting discrepancies is not heavily skewed, and that many respondents make only one or two mistakes instead of being inaccurate for the majority of the transitions they were asked to remember.

While informative, as our study is the first to study the accuracy of LEC information on important life-course transitions over the entire life span in an offender sample, the nature of the sample also limits the generalizability of the current findings. First, the age range of our sample, and the extended period of recall may have negatively affected the accuracy of especially the timing of self-reported events. Given that most marriages and births of children tend to occur between the ages of twenty and thirty-five and given that the average age of the respondents at the time of the interview was 60, respondents were de facto asked to recall events that happened quite some time ago. The age range of our sample also makes that marriages, divorces, and childbirths respondents were asked to recall in majority took place in the late nineteen-seventies and early eighties. While becoming a parent is arguably still an equally salient life event in the lives of modern day parents, this is not necessarily so for marriage and especially divorce. During the time period our respondents experienced most of their divorces, divorce was far less common in Dutch society than it is today. Divorce being a much rarer event, the personal and societal reactions to a person divorcing may have been more outspoken during this period, increasing the saliency of the event. Analogously, cohabitation before marriage has increased over the last decades, arguably making the transition to marriage fuzzier in later generations. Results from this study therefore do not generalize to younger cohorts.

The current study did not address the accuracy of LEC reported convictions or imprisonment. Instead of focusing on what are often the "dependent"

variables in life-course criminological research, here we focused on its "independent" variables, and aimed to assess the accuracy of LEC data on events thought to influence the likelihood of crime and conviction. The studies by Roberts and Wells (2010) and Morris and Slocum (2010), which we discussed in the introduction, covered a much shorter recall period than the current study, suggests that the LEC method may have limited value in recalling the timing of much repeated events such as arrests. Historical official records on arrests, convictions, and imprisonment are also often more readily available than data on for example cohabitation, employment and addiction. Assessing the accuracy of these and similar life-course transitions is therefore highly relevant, especially given these transitions' central role in life-course theories explaining desistance and persistence in crime.

By comparing data gathered by LEC with official registries, this study does not speak on the way LECs compare to ordinary survey methods in terms of reporting accuracy. Completing an LEC is more time consuming than filling out a regular survey and requires additional effort in the training of interviewers. While the use of LECs may have other benefits besides gathering more accurate data—such as limiting sample attrition and survey fatigue—that may justify its use, comparing LEC data to self-reported data gathered by means of ordinary surveys therefore remains an important avenue for future research in life-course criminology.

Finally, the current study compared self-reported information gathered using the LEC with information present in official registries. Whether these official data themselves accurately represent reality necessarily remained unaddressed. To the extent that concordance of LEC and official data speaks on the validity of the information gathered, official data on some life events, such as marriage and childbirth, can be assumed to be fairly accurate, also given the incentives inherent to reporting such events, like eligibility for receiving child benefits. Other officially recorded transitions, like divorce, are likely to accurately represent reality to a far less extent. Analogous to the remark that it may not be marriage a such that promotes desistance, but rather the quality of the bond represented by the act of marriage (Laub, Nagin, & Sampson, 1998), researchers relying solely on official data regarding life events and transitions should be conscious of the fact that official registrations—like the date that legally marks a divorce—do not always neatly coincide with the qualitative changes in life circumstances—such as a change in the experienced level of social control—that are theorized to underlie these events effect on criminal development. Triangulation of data sources therefore remains pivotal, especially for offender samples whose disorganized lives take place in the peripheral vision of national and local authorities. Results of the current analyses provide the empirical underpinning for using the LEC as a promising method in this regard.

Funding

The data collection for the Criminal Career and Life-course Study was accommodated by Investment Grant NWO Medium [grant number 480-09-010].

Disclosure Statement

No potential conflict of interest was reported by the authors.

ORCID

Nikki van Gerwen http://orcid.org/0000-0003-4407-256X

References

Auriat, N. (1993). "My wife knows best" : A comparison of event dating accuracy between the wife, the husband, the couple, and the belgium population register. *Public Opinion Quarterly, 57*(2), 165-190.

Axinn, W. G., Pearce, L. D., & Ghimire, D. (1999). Innovations in life history calendar applications. *Social Science Research, 28*, 243-264.

Bahr, H. M., & Houts, K. C. (1971). Can you trust a homeless man? A comparison of official records and interview responses by Bowery Men. *Public Opinion Quarterly, 35* (3), 374-382.

Belli, R. F. (1988). 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. doi:10.1080/741942610

Belli, R. F., Shay, W. L., & Stafford, F. P. (2001). Event history calendars and question list surveys. *Public Opinion Quarterly, 65*(1), 45-74.

Belli, R. F., Smith, L. M., Andreski, P. M., & Agrawal, S. (2007). Methodological comparisons between CATI event history calendar and standardized conventional questionnaire instruments. *Public Opinion Quarterly, 71*(4), 603-622. doi:10.1093/poq/nfm045

Bellie, 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.

Bernasco, W., & Steenbeek, W. (2017). More places than crimes: Implications for evaluating the law of crime concentration at place. *Journal of Quantitative Criminology, 33*, 451-467. doi:10.1007/s10940-016-9324-7

Blokland, A. (2005). *Crime over the life span; trajetories of criminal behavior in Dutch offenders*. Leiden: NSCR.

Blokland, A., Nagin, D., & Nieuwbeerta, P. (2005). Life span offending trajactories of a Dutch conviction cohort. *Criminology, 43*(4), 919-954.

Blokland, A., & Nieuwbeerta, P. (2010). Life course criminology. In S. G. Shoham & P. Knepper (Eds.), *International handbook of criminology* (pp. 51-94). Boca Raton, FL: CRC Press/Taylor and Francis.

Brener, N. D., Billy, J. O. G., & Grady, W. R. (2003). Assessment of factors affecting the validity of self-reported health-risk behavior among adolescents: Evidence from the scientific literature. *Journal of Adolescent Health, 33*, 436-457. doi:10.1016/S1054-139X(03)00052-1

Caspi, A., Moffitt, T. E., Thornton, A., Freedman, D., Amell, J. W., Harrington, H., & 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*(2), 101-114.

Coughlin, S. S. (1990). Recall bias in epidemiologic studies. *Journal of Clinical Epidemiology, 43*(1), 87-91. doi:10.1016/0895-4356(90)90060-3

Eisenhower, D., Mathiowetz, N. A., & Morgenstein, D. (1991). Recall errors: Sources and bias reduction techniques. In P. P. Biemer, R. M. Groves, & L. E. Lyberg (Eds.), *Measurement errors in surveys* (pp. 127-144). New York: Wiley.

Elder, G. H. J. (1985). Perspectives on the life course. In G. H. J. Elder (Ed.), *Life course dynamics* (pp. 23-49). Ithaca, NY: Cornell University Press.

Elliott, D. S. (1982). Review Essay. *Criminology, 20*(3-4), 527-538. doi:10.1111/j.1745-9125.1982.tb00476.x

Engel, L. S., Keifer, M. C., & Zahm, S. H. (2001). Comparison of a traditional questionnaire with an icon/calendar-based questionnaire to assess occupational history. *American Journal of Industrial Medicine, 40*(5), 502-511.

Farrington, D. P. (1988). *Studying changes within individuals: The causes of offending. in studies of psychological risk: The power of longitudinal data*. New York, NY: Cambridge University Press.

Freedman, D., Thornton, A., Camburn, D., Alwin, D., & Young-DeMarco, L. (1988). The life history calendar: A technique for collecting retrospective data. *Sociologial Methodology, 18*, 37-68.

Hindelang, M. J., Hirschi, T., & Weis, J. G. (1981). *Measuring delinquency*. Beverly Hills, CA: Sage.

Hollin, C. R., & Palmer, E. J. (2006). Criminogenic need and women offenders: A critique of the literature. *Legal & Criminological Psychology, 11*, 179-195. doi:10.1348/135532505X57991

Horney, J. (2001). Criminal events and criminal careers: An integrative approach to the study of violence. In R. F. Meier, L. W. Kennedy, & V. F. Sacco (Eds.), *The process and structure of crime* (pp. 41-167). New Brunswick: Transaction.

Kimura, D., & Clarke, P. G. (2002). Women's advantage on verbal memory is not restricted to concrete words. *Psychological Reports, 91*, 1137-1142.

Laub, J. H., & Sampson, R. J. (2003). *Shared beginnings, divergent lives: Delinquent boys to age 70*. Cambridge, MA: Harvard University Press.

Laub, J. H., Nagin, D. S., & Sampson, R. J. (1998). Trajectories of change in criminal offending : Good marriages and the desistance process. *American Sociological Review, 63*(2), 225-238. Retrieved from http://www.jstor.org/stable/265732

Levine, B., Svoboda, E., Hay, J. F., Winocur, G., & Moscovitch, M. (2002). Aging and autobiographical memory: Dissociating episodic from semantic retrieval. *Psychology and Aging, 17*(4), 677-689. doi:10.1037/0882-7974.17.4.677

Lewis, D., & Mhlanga, B. (2001). A life of crime: The hidden truth about criminal activity. *International Journal of Market Research, 42*(2), 217-240.

Liberman, A. M. (Ed.). (2008). *The long view of crime: A synthesis of longitudinal research*. New York, NY: Springer. doi:10.1007/978-0-387-71165-2

Mackenzie, D. L., & Li, S. D. (2002). The impact of formal and informal social controls on the criminal activities of probationers. *Journal of Research in Crime and Delinquency, 39*(3), 243-276. doi:10.1177/002242780203900301

Mathiowetz, N. A., & Duncan, G. (1988). Out of work, out of mind: Response errors in retrospective reports of unemployment. *Journal of Business & Economic Statistics, 6*(2), 221-229. doi:10.1080/07350015.1988.10509656

Morris, N. A., & Slocum, L. A. (2010). The validity of self-reported prevalence, frequency, and timing of arrest: An evaluation of data collected using a life event cal-

endar. *Journal of Research in Crime and Delinquency, 10*(5), 1-31. doi:10.1177/0022427809357719

Porcellato, L., Carmichael, F., & Hulme, C. (2016). Using occupational history calendars to capture lengthy and complex working lives: A mixed method approach with older people. *International Journal of Social Research Methodology, 19*(3), 269-286. doi:10.1080/13645579.2014.988005

Roberts, J., & Horney, J. (2009). The life event calendar method in criminological research. In A. R. Piquero & D. Weisburd (Eds.), *Handbook of Quantitative Criminology* (pp. 189-321). New York, NY: Springer.

Roberts, J., & Wells, W. (2010). The validity of criminal justice contacts reported by inmates: A comparison of self-reported data with official prison records. *Journal of Criminal Justice, 38*(5), 1031-1037.

Schröder, M. (2011). *Retrospective data collection in the survey of health, ageing and retirement in Europa. SHARELIFE methodology*. Mannheim: MEA.

Schubert, C. A., Mulvey, E. P., Steinberg, L., Cauffman, E., Losoya, S. H., Hecker, T., & Knight, G. P. (2004). Operational lessons from the pathways to desistance project. *Youth Violence and Juvenile Justice, 2*(3), 237-255. doi:10.1177/1541204004265875

Skardhamar, T., & Savolainen, J. (2014). Changes in criminal offending around the time of job entry: A study of employment and desistance. *Criminology, 52*(2), 263-291. doi:10.1111/1745-9125.12037

Skowronski, J. J., & Thompson, C. P. (1990). Reconstructing the dates of personal events: Gender differences in accuracy. *Applied Cognitive Psychology, 4*(5), 371-381.

Slocum, L. A., Simpson, S. S., & Smith, D. A. (2005). Strained lives and crime: Examining intra-individual variation in strain and offending in a sample of incarcerated women. *Criminology, 43*, 1067-1110.

Stafford, F. P. (2009). Timeline data collection and analysis: Time diary and event history calender methods. In R. F. Bellie, F. P. Stafford, & D. F. Alwin (Eds.), *Calendar and time diary* (pp. 13-30). Thousand Oaks, CA: Sage.

Sudman, S., & Bradburn, N. M. (1973). Effects of time and memory factors on response in surveys. *Journal of American Statistical Association, 68*(344), 805-815.

Sutton, J. E. (2010). A review of the life-events calendar method for criminological research. *Journal of Criminal Justice, 38*(5), 1038-1044. doi:10.1016/j.jcrimjus.2010.07.006

Sutton, J. E., Bellair, P. E., Kowalski, B. R., Light, R., & Hutcherson, D. T. (2011). Reliability and validity of prisoner self-reports gathered using the life event calendar method. *Journal of Quantitative Criminology, 27*(2), 151-171. doi:10.1007/s10940-010-9101-y

Thornberry, T. P., & Krohn, M. D. (2000). The self-report method for measuring delinquency and crime. In R. D. Duffee, R. D. Crutchfield, S. Mastrofski, L. Marzerolle, D. McDowall, & B. Ostrom (Eds.), *Measurement and analysis of crime and justice* (pp. 33—83). Washington, DC: Department of Justice.

Uggen, C., & Thompson, M. (2003). The socioeconomic determinants of ill-gotten gains: Within-person changes in drug use and illegal earnings. *American Journal of Sociology, 109*(1), 146-185. doi:10.1086/378036

van de Weijer, S., Thornberry, T., Bijleveld, C., & Blokland, A. (2015). The effects of parental divorce on the intergenerational transmission of crime. *Societies, 5*, 89-108. doi:10.3390/soc5010089

Wheaton, B., & Gotlib, I. H. (1997). Trajectories and turning points over the life course: Concepts and themes. In I. H. Gotlib & B. Wheaton (Eds.), *Stress and adversity over the life course: Trajectories and turning points* (pp. 1—25). New York, NY: Cambridge University Press.

Wittebrood, K., & Nieuwbeerta, P. (2000). Criminal victimization during one's life course: The effects of previous victimization and patterns of rounte activities. *Journal of Research in Crime and Delinquency, 37*(1), 91-122.