2018

Memory Gaps in the American Time Use Survey. Investigating the Role of Retrieval Cues and Respondents' Level of Effort

Antje Kirchner
RTI International, akirchner@rti.org

Ana Lucía Córdova-Cazar
Universidad San Francisco de Quito Quito

Robert F. Belli
University of Nebraska-Lincoln Lincoln

Robert F. Belli
University of Nebraska - Lincoln, bbelli2@unl.edu

Survey Research Methods (2018)
Vol. 12, No. 3, pp. 231-245
doi:10.18148/srm/2018.v12i3.7257

ISSN 1864-3361
http://www.surveymethods.org

Memory Gaps in the American Time Use Survey. Investigating the Role of Retrieval Cues and Respondents' Level of Effort

Antje Kirchner
RTI International
Durham, USA

Robert F. Belli
University of Nebraska-Lincoln
Lincoln, USA

Ana Lucía Córdova-Cazar
Universidad San Francisco de Quito
Quito, Ecuador

Caitlin E. Deal
American Institutes for Research
Washington D.C., USA

Unaccounted respondent memory gaps – i.e., those activity gaps that are attributed by interviewers to respondents' memory failure – have serious implications for data quality. We contribute to the existing literature by investigating interviewing dynamics using paradata, distinguishing temporary memory gaps, which can be resolved during the interview, from enduring memory gaps, which cannot be resolved. We investigate factors that are associated with both kinds of memory gaps and how different response strategies are associated with data quality. We investigate two hypotheses that are associated with temporary and enduring memory gaps. The motivated cuing hypothesis posits that respondents who display more behaviors related to the presence and use of retrieval cues throughout the survey will resolve temporary memory gaps more successfully compared to respondents displaying fewer such behaviors. This should result in overall lower levels of enduring memory gaps. The lack of effort hypothesis suggests that respondent who are less eager to participate in the survey will expend less cognitive effort to resolve temporary memory gaps compared to more motivated respondents. This should then result in a positive association with enduring memory gaps and no association with temporary memory gaps. Using survey and paradata from the 2010 ATUS, our analyses indicate that, as hypothesized, behaviors indicating the use of retrieval cues are positively associated with temporary memory gaps and negatively associated with enduring memory gaps. Motivated respondents experiencing memory difficulties overcome what otherwise would result in enduring memory gaps more successfully compared to other respondents. Indicators of lack of effort, such as whether or not the respondent initially refused to participate in the survey, are positively associated with enduring memory gaps suggesting that reluctant respondents do not resolve memory gaps. The paper concludes with a discussion of implications for survey research.

Keywords: American Time Use Survey, memory gaps, paradata, response strategy

1 Introduction

When we are asked to remember an event, our memory is typically dependent on information from two components (Tulving, 1974, p. 74). According to Tulving, the first results from experiencing an event and is referred to as the memory trace or encoding of the event in a person's memory storage. The second is the retrieval cue that is generally understood as information present in the person's cognitive environment at the time of retrieval. Remembering only occurs when both components are present. Past research on memory and forgetting suggests "when we forget something we once knew, it does not necessarily mean that the memory trace has been lost" (Tulving, 1974, p. 74). Instead, research stresses the importance of effective retrieval cues that will oftentimes lead to successful remembering of information that would otherwise be inaccessible (Al Baghal, Belli, Phillips, & Ruther, 2014; Belli, Lee, Stafford, & Chou, 2004; Mulligan & Picklesimer, 2012).

Using these insights of forgetting as a cue-dependent phenomenon, survey researchers have used event history calendars and time-use diaries that overcome some of the difficulties associated with recall of retrospective (autobiographical) events. These forms of surveying are based upon a flexible conversational exchange between interviewers and respondents (Belli & Callegaro, 2009; Stafford & Belli, 2009). By

Contact information: Antje Kirchner, RTI International, 3040 E. Cornwallis Road, RTP, NC 27709: akirchner@rti.org)

facilitating a more optimal use of retrieval behaviors and cuing of past events in respondents' autobiographical pasts, calendar and diary interviews have reliably led to better data quality of retrospective reports compared to conventional standardized interviewing; see also Belli (2014), Belli, Bilgen, and Al Baghal (2013), Belli, Lepkowski, and Kabeto (2001), Bilgen and Belli (2010). The main assumption of these methods is, however, that respondents are willing to expend the cognitive effort to recall retrospective events and when offered use such cues[1] – an assumption that has not previously been tested and which we aim to investigate more closely.

More specifically, the American Time Use Survey (ATUS) is one example of time-use surveys asking respondents to report everything they did the day prior to the interview in a chronological order.[2] ATUS, a nationally representative survey of the US population, is conducted by the US Census Bureau for the Bureau of Labor Statistics and is designed to provide high-quality survey data on how Americans spend their time. Although data quality aspirations are high in the ATUS (Bureau of the Census-United States Department of Commerce & Bureau of Labor Statistics- and United States Department of Labor, 2011; Horrigan & Herz, 2004), and ATUS data have been used increasingly by academics, governmental agencies and the press,[3] in the 2010 ATUS, 611 (4.63 percent) respondents had at least one reported memory gap. Memory gaps are time periods in which respondents cannot remember what they did. These activity gaps that are attributed to memory failure by interviewers are then coded as "Gap/Can't remember." For respondents with memory gaps, these periods account for an average of 1 hour and 10 minutes per day. It is plausible to assume that these numbers are a conservative estimate masking unreported periods of forgetting, that is, when respondents do not know or do not admit having forgotten an activity or when respondents claim they forgot what it is they did (e. g., motivated underreporting, Eckman et al., 2014). Furthermore, given the short reference period of reporting on the previous 24 hours that occurred the day before the interview in the ATUS, this issue will likely be exacerbated for time use surveys with longer reference periods.

Although some research has investigated memory gaps in time use surveys (e. g., Al Baghal et al., 2014; Belli et al., 2013), the question remains as to whether these memory gaps are purely the result of forgetting or whether respondents actually engage in some form of satisficing by indicating a misreport or a "don't know" when they would actually be able to remember the forgotten activity. If the former assumption holds, this would imply that retrieval cues could potentially help respondents who are willing to expend the cognitive effort to remember whereas the latter would suggest that retrieval cues will not necessarily be all that helpful. Likely, a mixture of both co-occurs.

To our knowledge, our study is the first to investigate whether and how respondents overcome memory gaps in time use surveys. While the use of retrieval cues and respondent effort leading up to a memory gap are often unobserved and occur internally, they can be inferred using paradata about the dynamics in the interview process leading up to a memory gap. This differentiation is possible by investigating response dynamics such as changes in responses and by differentiating between what we refer to as "temporary" and "enduring" memory gaps. We test these two mechanisms, that is, the motivation to use retrieval cues and respondent effort, leading up to an enduring memory gap using survey data and paradata from the 2010 ATUS.

The next section provides background information and derives testable hypotheses. In section 3, we provide an overview of the data and methods used. Section 4 presents the results of the analyses. Section 5 discusses the findings and the implications for future research.

2 Background and Hypotheses

Memory gaps in the ATUS activity data are a result of respondents reporting that they do not know or cannot remember what it is they did during a given time period. We differentiate *temporary memory gaps* in an activity that are defined by respondents initially reporting that they do not know or cannot remember what they did prior to forming a final response which can either be an *enduring memory gap* or a reported activity. *Enduring memory gaps* for a time period are defined as those memory gaps that were recorded by the ATUS interviewer as the final response, that is, they were not resolved throughout the interview, resulting in a don't know or can't remember activity in the final ATUS public release dataset.

Consider that respondents might initially not remember what they did after breakfast leading the interviewers to record "Gap/Can't remember" for those time periods (Bureau of the Census-United States Department of Commerce & Bureau of Labor Statistics- and United States Department of Labor, 2011, p. 53). ATUS interviewers are then trained to guide respondents through any memory lapse by using retrieval cues and help them to resolve these gaps by using conversational probing techniques (Phipps & Vernon, 2009). Such probes include visualization techniques, for example,

[1] Respondents could also potentially provide a random response instead of expending the cognitive effort. This behavior, however, would lead to follow-up questions the respondent would be unable to answer and hence seems implausible.

[2] For more information see, Bureau of the Census-United States Department of Commerce and Bureau of Labor Statistics- and United States Department of Labor (2011, p. 1), Hamermesh, Frazis, and Stewart (2005), Phipps and Vernon (2009).

[3] For an overview see http://www.bls.gov/tus/research.htm.

by asking respondents to picture where they were after breakfast, or whom they were with. If respondents engage with the interviewers in seeking to remember an activity, but still do not remember what they did, these temporary memory gaps will eventually result in final or enduring memory gaps in the ATUS dataset. If respondents, however, do remember the activity, that they went, for example, jogging after breakfast, the temporary memory gaps will not be recorded in the final data but result in a "normal" activity in the ATUS dataset. Finally, if respondents do not engage with the interviewers at all, they would then simply have an enduring memory gap without the initial presence of a temporary memory gap. By these definitions, 74.7% of respondents with temporary memory gaps in the ATUS 2010 resolve those into a reported activity whereas 25.3% resulted in enduring memory gaps. The majority of enduring memory gaps (90.0%) reported by respondents were not preceded by temporary memory gaps.

The above examples assume that respondents who state that they don't know or can't remember an activity are not only able but also motivated to expend the cognitive effort to attempt to resolve this memory gap during the interview. Displaying this kind of effort and seek to use retrieval cues, whether effective or not, would be indicative of respondent motivation. Lack of such effort, on the other hand, could be the result of satisficing behavior that leads to misreporting or forgetting without effort to retrieve and either would result in respondents providing a "Gap/Can't remember" response, that is, resulting in only an enduring memory gap. Analyses focusing exclusively on the relationship between respondent effort or motivation and the final recorded memory gaps are unable to differentiate between different mechanisms leading up to an enduring memory gap. Furthermore, they are confounded to the extent that respondents with different levels of motivation might differ in terms of how well they remember activities to begin with. Hence, controlling for respondent characteristics and whether a temporary memory gap occurred is crucial to understanding whether respondents who report a memory gap are willing to expend the necessary cognitive effort to remember past events and can, at times, successfully overcome temporary memory gaps. To separate these mechanisms, we investigate the relationship between the motivation to use retrieval cues and indicators of respondent effort, and temporary and enduring memory gaps while controlling for respondent characteristics. To measure the motivation to use retrieval cues and differentiate other forms of respondent effort we investigate proxy indicators derived from interviewing dynamics.

2.1 Retrieval cues and memory gaps

Ideally, respondents effortfully perform four cognitive processes when providing responses in a survey. These steps include question comprehension, information retrieval, judgment and reporting of their final response (e. g., Tourangeau, 1984). To recall activities over the 24-hour period ATUS interviewers are trained in different cuing techniques to assist respondents at the information retrieval stage. These techniques are based on mechanisms that are known to structure autobiographical memory (e. g., Belli, 2014). With time diaries such as the ATUS, temporal or sequential cuing predominates and will emphasize temporal properties of autobiographical knowledge that exist within themes and use certain events or activities as anchors to assist remembering temporally adjacent activities (Al Baghal et al., 2014). For example, remembering and reporting driving one's spouse to her employment will assist in remembering that upon returning home, one engaged in yard work while alone.

Supplementing the likelihood of successful sequential retrieval, ATUS interviewers are trained to use specific retrieval techniques to assist respondents remember forgotten activities (Phipps & Vernon, 2009). Visualization as one such technique is based on the assumption that events are stored in an associative manner such that remembering where an activity took place and whom a respondent was with for that activity can help respondents remember what it was that they did and therefore increase data quality (Belli, 1998; Tourangeau, 2000). Wagenaar (1986), for example, found that such "where" cues enhance people's memories as does additional information about an activity (Stafford, 2009). Retrieval cues such as who the respondent was with during an activity classify as such additional information and should hence foster recall for different activities. Al Baghal et al's (2014) findings support the importance of associations among details in remembering activities: When such memory associations were deficient, such as when respondents report an activity when they are alone or being at an ordinary place such as one's home, there is a significantly greater likelihood of the reporting of memory gaps for the subsequent period. To summarize, remembering who one was with and where one was during a given activity and making adjustments to the details of activities may also assist in providing higher quality linkages between activities throughout the day and may activate the memory trace for otherwise forgotten activities.

A key challenge to studying these retrieval processes and the use of retrieval cues is that they are often unobserved and can only be inferred using variables indicative of interviewing dynamics and reporting behavior such as the changes in respondents' reports of whom the respondent was with or where the activity took place. For example, let us assume respondents express uncertainty about aspects of an activity, such as where or with whom it occurred. If this is the case, then the number of response changes to those details of the reported activities, including changes of whom the respondent was with and where the activity took place would be indicative of the presence of motivation on the part of respondents to seek to remember details of activities. In addition, interviewers should be responsive to respondent uncertainty

by engaging in more probing using different cuing techniques such as, for example, asking respondents what activities took place where and with whom either earlier or later in time, ideally leading respondents to provide more detailed information or alternative changed information.

Another possibility, however, is that respondents who are changing their responses may be doing so in a haphazard way, perhaps because of difficulty in remembering or due to a lack of motivation (Heerwegh, 2003; Krosnick, 1991). A change in responses would then not signify the presence and use of retrieval cues and an increase in data quality but rather the opposite.

If the former assumption holds, that is, that changes to responses in activity details can be used as proxies for the motivation to remember or the presence of retrieval cues, then in turn the *motivated cuing hypothesis* would suggest that more frequent response changes should be associated with a lower number of enduring memory gaps and a higher number of respondents who reported a temporary memory gap (since otherwise cueing would not be necessary). If the latter assumption holds, we would observe the contrary, namely, that changes in responses are unrelated to enduring memory gaps. This assumption would either suggest that the changes in who and where do not capture the presence and use of retrieval cues or that retrieval cues are overwhelmingly unsuccessful in overcoming memory gaps.

2.2 Respondent motivation and memory gaps

The use of retrieval cues and strategies will, however, only lead to a reduction in enduring memory gaps if respondents are motivated to expend the cognitive effort to recall forgotten activities. To do so, ideally, respondents would effortfully and ably perform all four cognitive processes. Krosnick (1991) refers to the effortful and able processing of each phase as optimizing behavior. Assuming that memory traces can be retrieved using retrieval cues, respondents should then be able to resolve temporary memory gaps as activities for those time periods, provided they are willing and able to expend the additional cognitive effort, ultimately resulting in a greater number of reported activities and fewer enduring memory gaps. In contrast, satisficing behavior occurs whenever some of these phases are skipped, or they are processed with less cognitive effort or ability than necessary. This behavior would manifest in merely satisfactory responses and can result in incomplete, biased, or in its strongest form, missing data. Alternatively, respondents might also choose to misreport (Tourangeau, Kreuter, & Eckman, 2015) and claim they forgot an activity for reasons such as social desirability. For example, respondents might not want to reveal to interviewers that they were involved in sexual activities during those time periods in question. In these instances, enduring memory gaps can be the result not only of forgetting but also of low respondent motivation resulting in satisficing behaviors or misreporting that occurs despite attempts by interviewers to provide memory cues to the respondents.

Interviewing dynamics can indicate reluctance and lack of respondent effort resulting in misreporting and satisficing behavior (e. g., Fricker & Tourangeau, 2010). For example, respondents with a lower propensity to respond may be less engaged in forming a response or more generally provide lower quality data and speed through the interview (Groves & Couper, 1998; Holbrook, Green, & Krosnick, 2003; Kreuter, Müller, & Trappmann, 2010; Olson, 2006, 2013). Examples of behaviors of less engaged respondents could include answering before the interviewer completes the question, shortcutting definitions provided by the interviewer, or a decreased likelihood to ask for clarification. Less cooperative respondents might also engage in other satisficing behavior such as straightlining, providing "don't know' responses or acquiescent responses (but see Kaminska, McCutcheon, & Billiet, 2010). Applied to the context of the ATUS, respondents with lower response propensities, that is, respondents who are harder to contact or more reluctant to participate in a study may be more prone to provide lower quality data by expending less cognitive effort to resolve memory gaps compared to more motivated respondents (e. g., Al Baghal et al., 2014; Fricker & Tourangeau, 2010; Krosnick, 1991; Olson, 2013). Hence, such lack of motivation would result in either not having a temporary memory gap to begin with and only having enduring memory gaps as temporary gaps may be indicative of an attempt to remember; or as the occurrence of a temporary memory gap that leads to an enduring memory gap as a temporary memory gap merely indicates an expression of uncertainty and there is insufficient motivation to use available cues to resolve the gap. According to *the lack of effort hypothesis*, a lack of effort would result in more frequent reports of enduring memory gaps and there may or may not be an association between the level of respondent effort and temporary memory gaps, depending on whether temporary memory gaps are an indication of an attempt to remember or not.

3 Data and Methods

3.1 Data description

We use survey data and paradata from the 2010 American Time Use Survey (ATUS; Bureau of Labour Statistics, 2010). The ATUS sample is drawn from households, who completed their eighth Current Population Survey (CPS) interview using a three-stage, stratified sampling procedure. Each year, a sample of around 24,000 households is drawn with an oversample of households who have children, Hispanic and black households. One person, aged 15 or older, from each sampled household is randomly selected and assigned a specific day of the week about which to report. Data for the ATUS are recorded using computer-assisted telephone interview-

ing. Data collection is continuous and occurs almost every day of the calendar year with an average completion time of approximately 20 minutes.

ATUS respondents are asked to report events in a 24-hour period starting the previous day at 4 a.m. until 4 a.m. on the interview day. Interviewers are trained specifically to use a set of scripted questions in combination with conversational interviewing to obtain information about these events in a chronological order. Interviewers record activities in predefined codes or verbatim responses, the time and details of the activities including individuals who were present during the activity and where the activity occurred. An exception to these details being asked for is when a respondent reports personal activities such as sleeping or grooming. For these, "whom" is not asked, as is the case when a respondent reports working or high school activities (Bureau of the Census-United States Department of Commerce & Bureau of Labor Statistics- and United States Department of Labor, 2011).

We use public release data from the 2010 ATUS in combination with paradata collected via Blaise, often referred to as audit trails. In these audit trails the Blaise software automatically records interviewers' keystrokes, that is, it captures detailed information about the interviewing process including substantive responses (Couper, Hansen, & Sadosky, 1997).[4] The response rate for the 2010 ATUS is 56.9 percent (AAPOR RR2) with 13,260 completes. We have paradata for 14,027 ATUS cases of which 13,193 could be merged to the ATUS data (note: the ATUS public release data exclude respondents with fewer than 5 activities while the paradata do not).[5] A total of 69 interviewers completed an average of 196.8 interviews ranging from 1 to 780 interviews. The mean completion time of the time-diary portion of the ATUS based on the paradata is 12.2 minutes ranging from 2.0 to 29 minutes (censored at the 1st and 95th percentile).

3.2 Measuring Memory Gaps

To assess whether or not respondents overcome temporary memory gaps we need to combine survey and paradata. Temporary memory gaps are derived from the 2010 ATUS paradata and measure whether a respondent reported any memory gap in an activity prior to her final response. The ATUS paradata logs any interviewer actions which can then be parsed into a structured dataset. For example, an interviewer might add an activity (action 1: "Gap/Can't remember") followed by adding information about the next activity (action 2: "Watching TV"), the duration / time (action 3: "1 hour"), whom the respondent was with (action 4: "Alone") or where the respondent was (action 5: "Respondent's home or yard"). The next action could be a change to the information about the previously forgotten activity (action 6: "Grooming") due to some retrieval cue provided by the interviewer or self-generated by the respondent. Coding this example would lead to action 1 being coded as a temporary memory gap and action 6 being coded as an activity, that is, grooming. 1.83% of all respondents had at least one temporary memory gap prior to their final reported activity. Given the bimodal distribution of temporary memory gaps we created an indicator variable (0 = no temporary memory gap, 1 = temporary memory gap). Enduring memory gaps are derived from the 2010 ATUS public release data and operationalized as whether or not respondents had at least one final report of "Gap/Can't remember" in their reported activities (0 = no enduring memory gap, 1 = enduring memory gap). 4.63% of all respondents report at least one enduring memory gap.

3.3 Independent Variables

We are primarily interested in whether respondents are motivated to use retrieval cues and are willing to expend the cognitive effort to seek to overcome temporary memory gaps or not. Since the *presence and use of retrieval cues* cannot be directly observed, we use proxies derived from paradata. The presence and use of motivated retrieval cues is proxied using the average number of changes in a response to "whom" the respondent was with or "where" the respondent was during a given activity. We calculated the average number of changes in those responses per respondent since both measures are highly correlated. This reduces the risk of multicollinearity in the subsequent analyses. The average number of changes in whom and where responses is 0.53 and ranges from 0 to 15.5. The majority of the respondents have no change. The distribution of the average number of changes in whom and where is displayed in Figure 1.

Respondent lack of effort is operationalized using indicators of reluctance to be interviewed from the ATUS contact histories and prior CPS data (see also, Fricker & Tourangeau, 2010): the number of contact attempts needed to obtain an interview, whether a respondent initially refused to participate (0=never refused, 1=initial refusal), information from the preceding CPS interview, that is, whether or not the household reference person refused to provide information on family income (0=valid response or don't know, 1=refusal, 2=no information), and whether or not the ATUS respondent is identical to the CPS reference person and henceforth previously participated in a related survey (0=same person, 1=dif-

[4]The audit trails will be publicly available at ICPSR as of 1-1-2019 (2010 American Time Use Survey CATI Paradata, ICPSR deposit number 37282). These audit trails were sanitized and do not contain any personally identifiable information. Public use files of the audit trails will be available at ICPSR as of 1-1-2019 (2010 American Time Use Survey CATI Paradata, ICPSR deposit number 37282).

[5]Overall 141 cases were coded as having insufficient data quality by the interviewers (or had a blank value). Excluding those cases from the analyses leaving an analytic sample size of 13,052 cases does not alter our substantive results.

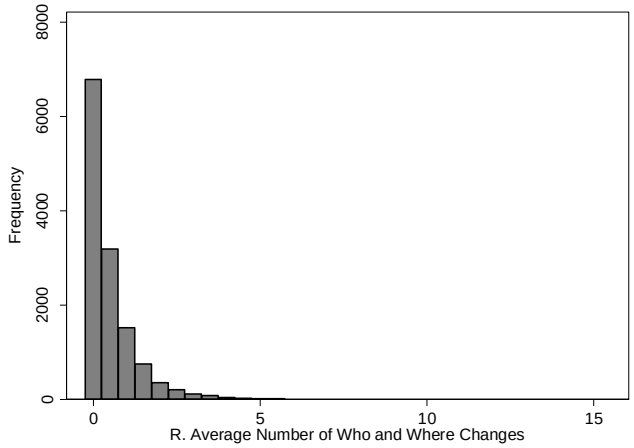

Figure 1. Average Number of "Whom" Changes and "Where" Changes

ferent person). The "no information" category in the family income is due to the fact that ATUS respondents drawn from "CPS households that completed their final interview in 2009 will have missing values for the edited family income variable" due to a change in measurement and imputation rules (Bureau of the Census-United States Department of Commerce & Bureau of Labor Statistics- and United States Department of Labor, 2011, p. 23).

It took an average number of 6.3 contact attempts, ranging from 0 to 52, to obtain a completed interview. Approximately 8 percent of all respondents refused to participate initially. Overall about 10 percent of the respondents refused to provide information on their family income. Almost 63 percent of all ATUS respondents are identical with the CPS respondent.

3.4 Control Variables

All models control for respondent characteristics including gender (0=Male, 1=Female), age (mean=46.64), education (0=High school or less, 1=Some college, associate degree, 2=Bachelor, 3=Graduate degree), race (0=White, 1=Nonwhite or multiple), ethnicity (0=Non-Hispanic, 1=Hispanic), employment status (0=Not in labor force, 1=Unemployed, 2=Employed), family income (0=Less than $30,000, 1=$30,000 to $74,999, 2=$75,000 and up), home ownership (0=Rent/Other, 1=Home owner), marital status (0=Married, 1=Other), presence of underage children in household (0=no, 1=yes), living in a metropolitan area (0=Metropolitan, 1=Non-metropolitan, 2=Not identified). Given that the probability of memory gaps might also be dependent on the number of reported activities (mean=19.43), the number of activities is included as an exposure variable or control (see section below).

All models account for interviewer within-study experience as operationalized using the number of completed interviews per interviewer (mean=196.77; Olson & Peytchev, 2007), to capture within-study experience, and the interviewer cooperation rate (mean=22.99%) as a second experience and quality measure. On one hand, interviewers with more experience and higher cooperation rates might differ systematically in their interviewing behavior (Olson, Kirchner, & Smyth, 2016) and either produce fewer enduring memory gaps or fail to record temporary memory gaps. The latter may be a result of experienced interviewers engaging in more probing before recording a (final) response. On the other hand, these interviewers might potentially recruit a different respondent pool that is less prone to report enduring memory gaps (West & Olson, 2010).

All continuous variables are centered at their grand mean. Table 1 presents an overview of the independent variables and controls including their operationalization and descriptive statistics.

4 Methods

We use two-level hierarchical logistic regression models to account for the clustering of respondents within interviewers (Hox, 1994; Hox, De Leeuw, & Kreft, 1991; Raudenbush & Bryk, 2002; Snijders & Bosker, 1999; Yan & Tourangeau, 2008). Table 2 shows the variability for each form of memory gap, that is, temporary and enduring memory gaps, across interviewers at the $p < 0.01$ level.

The between-interviewer variance or intraclass correlation (ICC) accounts for 14% of the total variance for temporary memory gaps (LR $\chi^2(1) = 39.91$; $p < 0.001$) and 19% of the total variance for the enduring memory gaps (LR $\chi^2(1) = 200.03$; $p < 0.001$) suggesting the appropriateness of using multilevel models.

To capture the interviewing dynamics resulting in a more complex relationship of temporary and enduring memory gaps we use multilevel structural equation modeling. Structural models for observed variables, allow the estimation of systems of equations and hence to model effects of mediating or intervening variables (Kline, 2011) such as temporary memory gaps. Temporary memory gaps have a dual role in these models: They are exogenous with respect to the response strategy, and they are endogenous with respect to enduring memory gaps and as such require a modeling strategy that allows for the simultaneous estimation of enduring memory gaps and temporary memory gaps. All models were estimated using Stata 15.1 GSEM and Figure 2 illustrates the estimated structural equation model using a path diagram[6].

The boxes in Figure 2 indicate observed variables (more specifically, for our independent variables a set of variables within each hypothesis) whereas circles contain unobserved,

[6]For the purpose of replication, all do-files are available in the online supplementary files.

Table 1
Descriptive Statistics of the Independent Variables (n = 13, 193)

Measure	Operationalization	Mean	SD	Min	Max
Motivated Cuing Hypothesis					
"Who" and "Where" changes	Avg. number of changes in "who" the respondent was with and "where" the respondent was	0.53	0.91	0	15.5
Lack of Effort Hypothesis					
Number of contact attempts	Number of contact attempts (call histories). Median split: 0=below median, 1=above median	6.32	6.87	0	52
Initial refusal	Whether a respondent ever refused to participate (call histories). 0=never refused, 1=initial refusal.	0.08	-	0	1
Item nonresponse	Family income item nonresponse (in CPS)				
	0=valid response or don't know,	0.86	-	0	1
	1=refusal,	0.10	-	0	1
	2=no information	0.04	-	0	1
CPS respondent	Whether or not the ATUS respondent is identical to the CPS respondent.				
	0=same, 1=different	0.37	-	0	1
Controls					
Female	Gender. 0=Male, 1=Female	0.56	-	0	1
Age	Age in years	46.64	17.74	15	85
Education	High school or less	0.42	-	0	1
	Some college, associate degree	0.27	-	0	1
	Bachelor	0.20	-	0	1
	Graduate degree	0.12	-	0	1
Race	0=White, 1=Nonwhite or Multiple	0.21	-	0	1
Ethnicity	0=Non-Hispanic, 1=Hispanic	0.14	-	0	1
Employment status	Not in labor force	0.35	-	0	1
	Unemployed	0.06	-	0	1
	Employed	0.59	-	0	1
Family income	Imputed:				
	Less than $30,000	0.23	-	0	1
	$30,000 to $74,999	0.28	-	0	1
	$75,000 and up	0.22	-	0	1
Home owndership	0=Rent/Other, 1=Home owner	0.72	-	0	1
Marital staus	0=Married, 1=Other	0.52	-	0	1
Presence of children	Presence of underage children in household (0=no, 1=yes)	0.47	-	0	1
Metropolitan area	Metropolitan	0.82	-	0	1
	Non-metropolitan	0.17	-	0	1
	Not identified	0.01	-	0	1
Activities	Number of reported activities	19.43	8.10	5	10
Interviewer Experience	Within-study experience Number of completed interviews	196.77	206.51	1	780
Cooperation rate	Interviewer cooperation rate	22.99		3.52	100

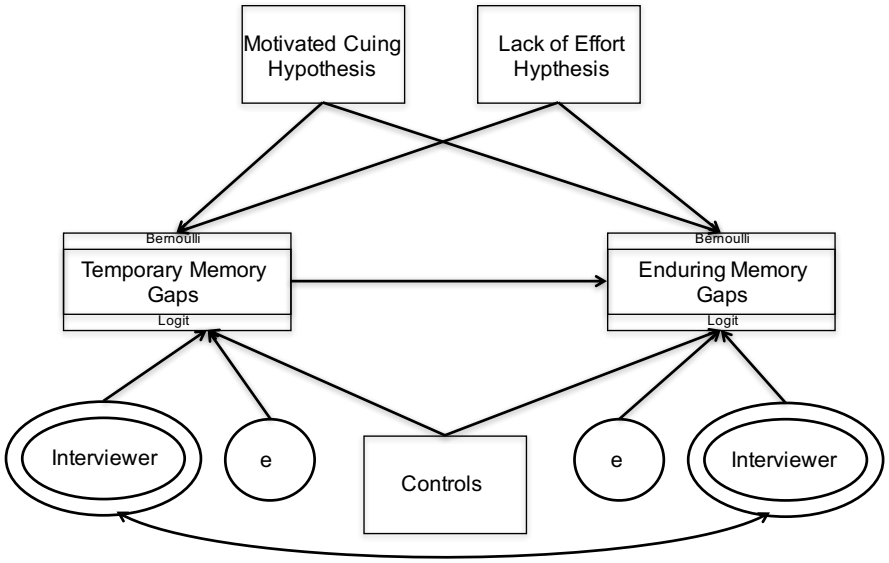

Figure 2. Underlying Structural Equation Model

Table 2
Interviewer Variance and Intraclass Correlations for Temporary and Enduring Memory Gaps

	Temporary memory gaps	Enduring memory gaps
Interviewer level variance	0.54**	0.78***
Intraclass correlation	0.14	0.19

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

or latent variables (Stata 2013). The double-ringed circles indicate the interviewer random effects. Each straight arrow indicates a linear equation and the curved arrow indicates that there is a covariance between the random effects.

Since memory gaps are also a rare event we retained all cases that displayed either a temporary or an enduring memory gap and subsampled ten percent of the remaining sample. This reduces the analytic sample size from n = 13,193 to n = 2,031[7].

5 Results

Table 3 provides the odds ratios, confidence intervals, and variance components modeling enduring memory gaps as a function of temporary memory gaps, motivated cuing, lack of respondent effort, and controls. We only display the coefficients for our main independent variables to maintain clarity. Results for the full models displaying all coefficients, including controls can be found in Appendix A1.

We only report statistically significant ($p < 0.05$) results in the text presenting exponentiated coefficients. Turning to the *Null Model* first, our results suggest that the probability of observing an enduring memory gap is significantly lower whenever an enduring memory gap is preceded by a temporary gap (OR = 0.60, $p < 0.01$).

The *Motivated Cuing Hypothesis* argues that respondents who display a higher presence and use of retrieval cues and who have a temporary memory gap will seek to resolve this memory gap, at times being successful at such resolution. Turning to the relationship between temporary memory gaps and retrieval cues we find that there is a positive association between the probability of a temporary memory gap and retrieval cues that is proxied using the number of changes in responses to whom the respondent was with or where the respondent was (OR = 1.20, $p < 0.01$). In line with our expectations, our results also suggest that a higher frequency of retrieval cues is in turn associated with fewer enduring memory gaps and an overall higher data quality (OR = 0.85, $p < 0.05$). Overall these findings support the Motivated Cueing Hypothesis.

Turning to the *Lack of Effort Hypothesis*, our results suggest that in line with our expectations, respondents who dis-

[7]With the exception of the relationship between temporary and enduring memory gaps, the substantive interpretation of the results is identical compared to using all available cases (effects tend to be somewhat more pronounced using all available cases). When using all available cases, the relationship between temporary memory gaps and enduring memory gaps is positive ($p < 0.001$) indicating that the presence of temporary memory gaps generally increases the probability of observing an enduring memory gap. When using the smaller analytic sample, this relationship is reversed suggesting that sample members with a temporary memory gap have an increased probability of resolving the enduring memory gap. This difference can be explained by the fact that the majority of the cases do not have any temporary or enduring memory gaps to begin with and retaining these cases in the model dominates the estimated relationship.

Table 3
*Estimated Odds Ratios, Confidence Intervals, and Variance Components for Hierarchical Models Predicting Memory Gap in Final and Temporary Response (*GSEM*)*

	Full Model		Combined Model	
	Odds Ratio	95% CI	Odds Ratio	95% CI
(a) ATUS enduring memory gap				
Temporary memory gap (ref. None)	0.60**	0.43, 0.83	0.48***	0.35, 0.68
Motivated Cueing Hypothesis				
Avg Number of Who and Where Changes	-	-	0.85*	0.75, 0.97
Lack of Effort Hypothesis				
Number of contact attempts	-	-	1.01	1.00, 1.03
Initial refusal	-	-	1.91***	1.31, 2.79
Family income refused (ref. valid entry)	-	-	0.92	0.65, 1.30
No information	-	-	0.97	0.58, 1.63
CPS respondent	-	-	0.87	0.68, 1.11
(b) Temporary memory gap				
Motivated Cueing Hypothesis				
Avg Number of Who and Where Changes	-	-	1.20**	1.05, 1.37
Lack of Effort Hypothesis				
Number of contact attempts	-	-	1.00	0.98, 1.02
Initial refusal	-	-	1.07	0.63, 1.82
Family income refused (ref. valid entry)	-	-	0.71	0.43, 1.16
No information	-	-	0.70	0.36, 1.38
CPS respondent	-	-	0.78	0.55, 1.11
Interviewer level variance: Enduring memory gap	0.70**		0.56**	
Interviewer level variance: Temporary memory gap	0.26+		0.18+	
Interviewer level covariance constant	0.35**		0.27*	
N	2,031		2,031	
AIC	3,845.59		3,733.01	

+ $p < .1$ * $p < .05$ ** $p < .01$ *** $p < .001$
Note: The combined model also includes controls. See Appendix A1 for full models.

play lower levels of effort and more reluctance are significantly more likely to have enduring memory gaps. While the number of contact attempts is not significantly related to enduring memory gaps (OR = 1.01, $p = 0.12$), whether or not the ATUS sample member initially refused to participate significantly increases the probability of observing an enduring memory gap (OR = 1.91, $p < 0.01$). Contrary to our expectations, respondents whose information regarding their family income is missing in the CPS are not more likely to exhibit enduring memory gaps compared to respondents with valid entries (OR = 0.92, $p = 0.63$). Whether or not the respondent previously participated in the CPS also does not affect their propensity to display enduring memory gaps (OR = 0.87, $p = 0.27$). These latter two null finding are not as surprising given that 37% of all ATUS respondents are not identical with the CPS respondent and both associations may be weakened as a result. In line with our expectations, none of the reluctance indicators is significantly associated with temporary memory gaps. These findings suggest that there is a certain amount of reduced level of effort among respondents with a lower response propensity to the extent that they do not try to resolve memory lapses. This lack of effort ultimately results in a higher frequency of enduring memory gaps for those respondents who were initially reluctant to participate in the survey. While the interviewer variance components are smaller in the Combined Model compared to the Null Model, they are still significantly different from zero. Figure 3 visualizes the results from the Combined Model; a blue dashed line indicates evidence for the motivated cuing hypothesis and the solid red line indicates evidence for the lack of effort hypothesis.[8]

[8]Both types of memory gaps are rare events potentially leading to biased parameter estimates. We also re-estimated all models using Firth's penalized maximum likelihood routine FIRTHLOGIT in

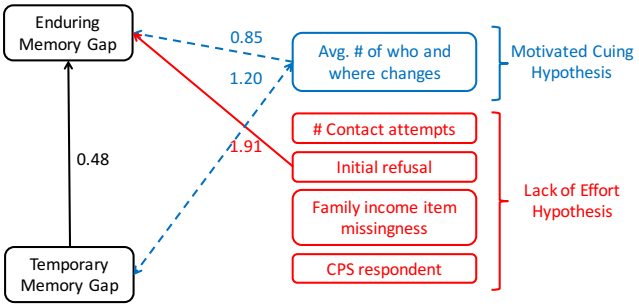

Figure 3. Estimated Odds Ratios for Hierarchical Models Predicting Temporary and Enduring Memory Gaps ($p < 0.05$)

The results for the controls can be found in Appendix A1. Consistent with the existing literature (e. g., Al Baghal et al., 2014; Knäuper, Belli, Hill, & Herzog, 1997) as age increases the number of enduring memory gaps increases (OR = 1.02 ,$p < 0.001$). Respondents with some college experience or associates degree (OR = 0.77, $p < 0.05$) display significantly fewer enduring memory gaps compared to respondents with a high school degree or less. Additionally, respondents interviewed by interviewers with a high cooperation rate have significantly fewer enduring memory gaps compared to those interviewed by interviewers with a low cooperation rate (OR = 0.96, $p < 0.01$). Finally, the higher the number of reported activities, the higher the probability of an enduring memory gap (OR = 1.05, $p < 0.001$). This finding is also consistent with the existing literature on complex autobiographical histories (Belli et al., 2013)).

Turning to temporary memory gaps we see that again older respondents (OR = 1.02, $p < 0.001$), and those with some college experience or an associate's degree (OR = 1.43, $p < 0.05$) and those with a Bachelors' degree (OR = 1.61, $p < 0.05$) have a significantly higher probability to report a temporary memory gap. Finally, respondents with an income of $75,000 and above have a significantly decreased probability of having a temporary memory gap (OR = 0.60, $p < 0.05$).

The effects of the control variables vary across the different stages of memory gaps. For example, individuals with an income of $75,000 and above have a decreased probability of temporary memory gaps which this does not hold with respect to enduring memory gaps. This result suggests that individuals in this income category are not as efficient in resolving temporary memory gaps resulting in similar rates of enduring memory gaps compared to individuals earning less.

6 Discussion and Conclusion

Our goal in this research is to understand the mechanisms leading to enduring memory gaps and hence a decrease in data quality. Using information about interviewing dynamics extracted from the 2010 ATUS paradata, we investigate whether the number of enduring memory gaps can be reduced if respondents display a higher frequency of presence and use of retrieval cues (*Motivated Cuing Hypothesis*). As hypothesized, our results suggest that a more frequent display of retrieval cues and memory activation is associated with respondents more successfully resolving temporary memory gaps which in turn also leads to significantly fewer enduring memory gaps. Not surprisingly, respondents displaying behaviors indicative of lack of effort and motivation, that is, a lower response propensity, do not engage in the cognitive processes to resolve memory gaps and instead have more enduring memory gaps (*Lack Effort Hypothesis*).

To summarize, our results indicate that enduring memory gaps can be significantly reduced, although there is evidence that not all respondents are willing to go through the effort to do so. This result not only suggests that respondents who expend the cognitive effort can resolve memory gaps but also that the use of retrieval cues might have aided respondents in recalling activities in these time periods.

The implications of our results are threefold. First, we demonstrate that the availability of paradata can help researchers to better understand response processes and hence potentially lead to an improvement of data quality. Collecting paradata, however, is yet by no means a standard practice. Given the utility of paradata in our analyses we recommend that this kind of data should be collected by default for time use surveys and event-history calendars. Second, training interviewers how to establish rapport, use retrieval cues, and motivate respondents to expend the extra cognitive effort to engage more fully in the response process to attempt to resolve temporary memory gaps, especially for those respondents with a lower response propensity, is particularly important. Third, our findings provide support that the number of changes in whom respondents were with and where they were located is associated with a lower probability of enduring memory gaps. This finding suggests that helping respondents create a context, that is, whom they were with

---

Stata (Firth, 1993). Coefficients estimated using penalized maximum likelihood have been shown to be unbiased in cases with small n and very few events (e. g., Leitgöb, 2013). This particular method has been shown to be superior compared to other approaches dealing with rare events (e. g., Leitgöb, 2013). Since this approach is not implemented in any of the available multilevel structural equation modeling packages in Stata, we re-estimated the combined models presented in the paper in simplified models including interviewer fixed effects. Our results show that regressing enduring memory gaps on temporary memory gaps (OR = 0.50, $p < 0.001$) and the average number of who and where changes (OR = 0.86, $p < 0.05$) remains highly significant (LR $\chi^2(2) = 23.95$, $p < 0.001$) as is a regression of temporary memory gaps on the average number of who and where changes (OR = 1.16, $p < 0.05$; LR $\chi^2(1) = 4.15$, $p < 0.05$).

and where they were during an activity, may be associated with a significant decrease in the number of reported enduring memory gaps. Training interviewers to effectively use retrieval techniques, such as visualization, to probe in case of memory lapses is hence of utmost importance.

Our study is also subject to some limitations. First, the kind of temporary and enduring memory gaps that we investigate consist of those gaps that the respondent reports and assumes that interviewers consistently record their presence. Oftentimes, however, respondents might not remember that they forgot an activity, the unknown unknowns. At other times, respondents might be aware of this forgetting but might not want to admit their forgetfulness to the interviewer and instead report something else as a form of motivated misreporting (see Tourangeau et al., 2015). Interviewers, on the other hand, might not record a (temporary) memory gap before engaging in more probing. In that instance, the paradata would never record a temporary memory gap but, depending on the success of the probing, would record an activity or an enduring memory gap. Both processes lead to an undercount of temporary and enduring memory gaps. Second, the ATUS public release data does not include data for respondents with fewer than five activities. This is particularly problematic as these respondents may represent a population with the poorest ability to remember. Third, a "Gap/Can't remember" response by a respondent might also mask implicit refusals, for example, due to concerns of social desirability. Without the use of appropriate mixture models, we are unable to differentiate memory gaps due to forgetting from memory gaps as a form of refusal. Fourth, there may also be an alternative explanation to our results. Respondents who resolve a temporary memory gap might actually be reacting to the interviewer. That is, they might want to appear as a "good" respondent which might result in respondents reporting a random activity although they do not actually remember what it is they did to please the interviewer. While this is a valid concern, our results in combination with the associations to the response changes suggest otherwise. Finally, the ATUS data does not contain any interviewer demographics or attitudes. This information could help to explain the significant interviewer variance components which suggest that interviewers use particular strategies when it comes to memory gaps and potentially shed light on specific, more successful interviewing strategies that lead to a higher data quality.

Acknowledgements

An earlier version of this paper was presented at the Annual Conference of the American Association for Public Opinion Research, May 2015, Hollywood, FL, and the Annual Conference of the Midwest Association for Public Opinion Research, Nov. 2016, Chicago, IL. This material is based upon work supported by the National Science Foundation under Grant No. SES-1132015. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

Al Baghal, T., Belli, R. F., Phillips, A. L., & Ruther, N. (2014). What are you doing now? Activity-level responses and recall failures in the American Time Use Survey. *Journal of Survey Statistics and Methodology*, *2*(4), 519–537.

Belli, R. F. (1998). The structure of autobiographical memory and the event history calendar: potential improvements in the quality of retrospective reports in surveys. *Memory*, *6*(4), 383–406.

Belli, R. F. (2014). Autobiographical memory dynamics in survey research. In T. J. Perfect & D. S. Lindsay (Eds.), *The SAGE handbook of applied memory* (pp. 366–384). Thousand Oaks, CA: Sage Publications.

Belli, R. F., Bilgen, I., & Al Baghal, T. (2013). Memory, communication, and data quality in calendar interviews. *Public Opinion Quarterly*, *77*(S1), 194–219.

Belli, R. F. & Callegaro, M. (2009). The emergence of calendar interviewing: a theoretical and empirical rationale. In R. F. Belli, F. P. Stafford, & D. Alwin (Eds.), *Calendar and time diary methods in life course research* (pp. 31–52). Thousand Oaks, CA: Sage.

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*(2), 185.

Belli, R. F., Lepkowski, J. M., & Kabeto, M. U. (2001). The respective roles of cognitive processing difficulty and conversational rapport on the accuracy of retrospective reports of doctor's office visits. In *Seventh conference on health survey research methods. DHHS publication no. (PHS) 01-1013. Hyattsvill, MD: U.S.* (p. 197).

Bilgen, I. & Belli, R. F. (2010). Comparison of verbal behaviors between calendar and standardized conventional questionnaires. *Journal of Official Statistics*, *26*(3), 481.

Bureau of Labour Statistics. (2010). American Time Use Survey. 2010 Microdata Files. Retrieved from https://www.bls.gov/tus/datafiles%5C_2010.htm

Bureau of the Census-United States Department of Commerce & Bureau of Labor Statistics- and United States Department of Labor. (2011). American Time Use Survey user's guide. Understanding ATUS 2003 to 2010. Retrieved from http://www.bls.gov/tus/atususersguide.pdf

Couper, M. P., Hansen, S. E., & Sadosky, S. A. (1997). Evaluating interviewer use of CAPI technology. In L. Lyberg, P. Biemer, M. Collins, E. de Leeuw, C. Dippo,

N. Schwarz, & D. Trewin (Eds.), *Survey measurement and process quality* (pp. 267–285). New York: Wiley and Sons.

Eckman, S., Kreuter, F., Kirchner, A., Jäckle, A., Tourangeau, R., & Presser, S. (2014). Assessing the mechanisms of misreporting to filter questions in surveys. *Public Opinion Quarterly*, *78*(3), 721–733.

Firth, D. (1993). Bias reduction of maximum likelihood estimates. *Biometrika*, *80*(1), 27–38.

Fricker, S. & Tourangeau, R. (2010). Examining the relationship between nonresponse propensity and data quality in two national household surveys. *Public Opinion Quarterly*, *74*(5), 934–955.

Groves, R. M. & Couper, M. P. (1998). *Nonresponse in household interview surveys*. John Wiley & Sons.

Hamermesh, D. S., Frazis, H., & Stewart, J. (2005). Data watch: the American Time Use Survey. *Journal of Economic Perspectives*, *19*(1), 221–232.

Heerwegh, D. (2003). Explaining response latencies and changing answers using client-side paradata from a web survey. *Social Science Computer Review*, *21*(3), 360–373.

Holbrook, A. L., Green, M. C., & Krosnick, J. A. (2003). Telephone versus face-to-face interviewing of national probability samples with long questionnaires: comparisons of respondent satisficing and social desirability response bias. *Public Opinion Quarterly*, *67*(1), 79–125.

Horrigan, M. & Herz, D. (2004). Planning, designing, and executing the BLS american time-use survey. *Monthly Lab. Rev. 127*, 3.

Hox, J. J. (1994). Hierarchical regression models for interviewer and respondent effects. *Sociological Methods & Research*, *22*(3), 300–318.

Hox, J. J., De Leeuw, E. D., & Kreft, G. G. (1991). The effect of interviewer and respondent characteristics on the quality of survey data: a multilevel model. In P. P. Biemer, R. M. Groves, L. E. Lyberg, N. Mathiowetz, & S. Sudman (Eds.), *Measurement errors in surveys* (pp. 439–461). New Xork: Wiley.

Kaminska, O., McCutcheon, A. L., & Billiet, J. (2010). Satisficing among reluctant respondents in a cross-national context. *Public Opinion Quarterly*, *74*(5), 956–984.

Kline, R. B. (2011). *Principles and practice of structural equation modeling*. Guilford Press.

Knäuper, B., Belli, R. F., Hill, D. H., & Herzog, A. R. (1997). Question difficulty and respondents' cognitive ability: the effect on data quality. *Journal of Official Statistics*, *13*, 181–199.

Kreuter, F., Müller, G., & Trappmann, M. (2010). Nonresponse and measurement error in employment research: making use of administrative data. *Public Opinion Quarterly*, *74*(5), 880–906.

Krosnick, J. A. (1991). Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied Cognitive Psychology*, *5*(3), 213–236.

Leitgöb, H. (2013). The problem of modeling rare events in ML-based logistic regression. assenssing potential remedies via MC simulations. Paper presented at the fifth Conference of the European Research Association, Ljubljana.

Mulligan, N. W. & Picklesimer, M. (2012). Levels of processing and the cue-dependent nature of recollection. *Journal of Memory and Language*, *66*(1), 79–92.

Olson, K. (2006). Survey participation, nonresponse bias, measurement error bias, and total bias. *International Journal of Public Opinion Quarterly*, *70*(5), 737–758.

Olson, K. (2013). Do non-response follow-ups improve or reduce data quality? A review of the existing literature. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *176*(1), 129–145.

Olson, K., Kirchner, A., & Smyth, J. (2016). Do interviewers with high cooperation rates behave differently? interviewer cooperation rates and interview behaviors. *Survey Practice*, *9*(2).

Olson, K. & Peytchev, A. (2007). Effect of interviewer experience on interview pace and interviewer attitudes. *Public Opinion Quarterly*, *71*(2), 273–286.

Phipps, P. A. & Vernon, M. K. (2009). Twenty-four hours: an overview of the recall diary method and data quality in the American Time Use Survey. In R. F. Belli, F. P. Stafford, & D. Alwin (Eds.), *Calendar and time diary methods in life course research* (pp. 109–128). Thousand Oaks, CA: Sage.

Raudenbush, S. W. & Bryk, A. S. (2002). *Hierarchical linear models: applications and data analysis methods*. Sage.

Snijders, T. A. B. & Bosker, R. J. (1999). *Multilevel analysis: an introduction to basic and advanced multilevel modeling*. London: Sage.

Stafford, F. P. (2009). Timeline data collection and analysis: time diary and event history calendar methods. In R. F. Belli & F. P. Stafford (Eds.), *Calendar and time diary methods in life course research* (pp. 13–30). Thousand Oaks, CA: Sage.

Stafford, F. P. & Belli, R. F. (2009). Future directions in calendar and time diary methods. In R. F. Belli & F. P. Stafford (Eds.), *Calendar and time diary methods in life course research* (pp. 13–30). Thousand Oaks, CA: Sage.

Tourangeau, R. (1984). Cognitive sciences and survey methods. In T. Jabine, M. Straf, R. Tanur, & R. Tourangeau (Eds.), *Cognitive aspects of survey methodology: building a bridge between disciplines* (pp. 73–100). Washington, DC: National Academy Press.

Tourangeau, R. (2000). Remembering what happened: memory errors and survey reports. In S. A. Arthur, J. S.

Turkkan, C. Bachrach, J. B. Jobe, H. S. Kurtzman, & V. S. Cain (Eds.), *The science of self-report: implications for research and practice* (pp. 29–47). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Tourangeau, R., Kreuter, F., & Eckman, S. (2015). Motivated misreporting: shaping answers to reduce survey burden. In U. Engel (Ed.), *Survey measurements: techniques, data quality, and sources of error* (pp. 22–41). Frankfurt/New York: Campus Verlag.

Tulving, E. (1974). Cue-dependent forgetting: when we forget something we once knew, it does not necessarily mean that the memory trace has been lost; it may only be inaccessible. *American Scientist*, *62*(1), 74–82.

Wagenaar, W. A. (1986). My memory: a study of autobiographical memory over six years. *Cognitive Psychology*, *18*(2), 225–252.

West, B. T. & Olson, K. (2010). How much of interviewer variance is really nonresponse error variance? *Public Opinion Quarterly*, *74*(5), 1004–1026.

Yan, T. & Tourangeau, R. (2008). Fast times and easy questions: the effects of age, experience and question complexity on web survey response times. *Applied Cognitive Psychology*, *22*(1), 51–68.

Appendix
Tables

(*Appendix table follows on next page*)

Table A1
Estimated Odds Ratios, confidence Intervals, and Variance Components for Hierarchical Models Predicting Memory Gap in Final and Temporary Response

	Full Model		Combined Model	
	Odds Ratio	95% CI	Odds Ratio	95% CI
(a) ATUS enduring memory gap				
Temporary memory gap (ref. None)	0.60**	0.43, 0.83	0.48***	0.35, 0.68
Motivated Cueing Hypothesis				
Avg Number of Who and Where Changes	-	-	0.85*	0.75, 0.97
Lack of Effort Hypothesis				
Number of contact attempts	-	-	1.01	1.00, 1.03
Initial refusal	-	-	1.91***	1.31, 2.79
Family income (ref. valid entry)				
Refused	-	-	0.92	0.65, 1.30
No information	-	-	0.97	0.58, 1.63
CPS respondent	-	-	0.87	0.68, 1.11
Controls				
Female (ref. Male)	-	-	0.95	0.75, 1.18
Age in years (centered)	-	-	1.02***	1.01, 1.03
Education (ref. High school or less)				
Some college; Associate degree	-	-	0.77*	0.59, 1.00
Bachelor	-	-	0.76+	0.56, 1.04
Graduate	-	-	0.80	0.53, 1.22
Race (ref. White)	-	-	0.94	0.71, 1.22
Hispanic	-	-	0.95	0.67, 1.35
Employment status (ref. Not in labor force)				
Unemployed	-	-	1.08	0.67, 1.75
Employed	-	-	0.82	0.63, 1.10
Family income (ref. $30,000 and less)				
$30,000 to $74,999	-	-	0.84	0.63, 1.10
$75,000 and up	-	-	0.91	0.64, 1.29
Home owner	-	-	1.12	0.85, 1.47
Marital status (ref. Married)	-	-	0.99	0.77, 1.27
Presence of child in hh <18	-	-	1.15	0.87, 1.53
Urbanicity (ref. Metropolitan)				
Non-metropolitan	-	-	1.00	0.76, 1.32
Not identified	-	-	1.04	0.32, 3.34
Number of reported activities	-	-	1.05***	1.03, 1.06
Interviewer Experience	-	-	1.00	1.00, 1.00
Interviewer Cooperation rate	-	-	0.96**	0.94, 0.99

Continues on next page

Continued from last page

	Full Model		Combined Model	
	Odds Ratio	95% CI	Odds Ratio	95% CI
(b) Temporary memory gap				
Motivated Cueing Hypothesis				
Avg Number of Who and Where Changes	-	-	1.20**	1.05, 1.37
Lack of Effort Hypothesis				
Number of contact attempts	-	-	1.00	0.98, 1.02
Initial refusal	-	-	1.07	0.63, 1.82
Family income (ref. valid entry)				
Refused	-	-	0.71	0.43, 1.16
No information	-	-	0.70	0.36, 1.38
CPS respondent	-	-	0.78	0.55, 1.11
Controls				
Female (ref. Male)	-	-	1.28	0.94, 1.74
Age in years (centered)	-	-	1.02***	1.01, 1.03
Education (ref. High school or less)				
Some college; Associate degree	-	-	1.43*	1.02, 2.02
Bachelor	-	-	1.61*	1.06, 2.42
Graduate	-	-	1.02	0.55, 1.90
Race (ref. White)	-	-	0.94	0.65, 1.34
Hispanic	-	-	0.82	0.50, 1.33
Employment status (ref. Not in labor force)				
Unemployed	-	-	0.83	0.42, 1.66
Employed	-	-	0.75+	0.53, 1.05
Family income (ref. $30,000 and less)				
$30,000 to $74,999	-	-	0.79	0.55, 1.14
$75,000 and up	-	-	0.60*	0.37, 0.97
Home owner	-	-	0.93	0.65, 1.32
Marital status (ref. Married)	-	-	0.99	0.71, 1.38
Presence of child in hh <18	-	-	1.18	0.79, 1.74
Urbanicity (ref. Metropolitan)				
Non-metropolitan	-	-	1.11	0.77, 1.59
Not identified	-	-	0.42	0.05, 3.42
Number of reported activities	-	-	1.01	0.99, 1.03
Interviewer Experience	-	-	1.00	1.00, 1.00
Interviewer Cooperation rate	-	-	0.99**	0.96, 1.03
Interviewer level variance: Enduring memory gap	0.70**		0.56**	
Interviewer level variance: Temporary memory gap	0.26+		0.18+	
Interviewer level covariance constant	0.35**		0.27*	
N	2,031		2,031	
AIC	3,845.59		3,733.01	

+ $p < .1$ * $p < .05$ ** $p < .01$ *** $p < .001$