Supplemental Materials for Accuracy and Completeness of Autobiographical Memory: Evidence from a Wearable Camera Study

Jason R. Finley and William F. Brewer

Contents:

Further Details on Daily Diary Studies

The first autobiographical memory studies to adopt a planned record approach relied on long-running systematic daily diaries made for the express purpose of studying autobiographical memory (Linton, 1975, 1978, 1982; Wagenaar, 1986; White, 1982, 1989, 2002). In all of these studies the researchers themselves were the sole participants. The researchers/participants in these heroic studies made daily diary records for acquisition periods of 1 to 6 years, and they used those records to test aspects of their autobiographical memory after retention intervals ranging from one day to 20 years. Each diary entry (typically one or two per day) constituted an "event" from that day, a specific episode selected by the researcher/participant to be distinct (Linton and Wagenaar) or selected haphazardly (White). Each researcher/participant followed a self-imposed recording protocol that included for example: a date, a description, and ratings along various scales (e.g., distinctiveness, emotionality, personal importance). Memory test tasks (i.e., dependent measures) included temporally ordering two randomly selected event records (Linton), self-rating of recollection when cued by part or all of a record (White), and recall of some aspects of a record when cued with other aspects (Wagenaar).

In a series of later diary studies, Thompson et al. (1996) recruited numerous undergraduates to record daily diaries for periods of 3 months (and several for up to 2.5 years), with instructions to record one event per day and to provide a variety of events (not just the most memorable). Testing sessions were conducted at the end of the recording period, with retention intervals thus ranging from 1 day to 2.5 years. Dependent measures included a self-reported rating of how well the participant remembered an event (1-7, where 1 was "not at all" and 7 was "perfectly"), recall of location and people involved in the event, and an estimation of the date of the event. Barclay and Wellman (1986) also conducted a daily diary study with several graduate students as participants, an acquisition period of 4 months of memorable daily events, and retention intervals up to 2.5 years. That study used a recognition test with several types of foils.

Although such diary methods do enable some objective evaluation of memory for personal experiences (see also Conway, 1990 for a review), they also have several disadvantages. First, there is bias in the selection of which experiences to record in the diary. Second, the very act of making the diary records changes participants' experiences and potentially alters their memories (see Brewer, 1988, p. 82; testing effect aka retrieval practice, Roediger & Butler, 2011; production effect, MacLeod & Bodner, 2017). That said, Thompson's 1982 "roommate study," which compared recorded events occurring in the life of the participant and events occurring in the life of the participant's (unaware) roommate, suggested minimal memory effects due to writing down the events.

Analysis of the Three Different Types of Omissions (Reported After Pictures)

During the self-scoring phase, once participants had seen their pictures and still had the pictures on the screen, they could report information that had been omitted in initial recall, and classify these omissions as: neglected, reminded, or forgotten. The instructions labeled these response types as follows: (Neglected) *I DID remember this in Free Recall but didn't write down*, (Reminded) *I DID NOT remember this in Free Recall and now I DO remember it*, (Forgotten) *I DID NOT remember this in Free Recall and I STILL DO NOT remember it.* Table 2to3 in the main text shows the ratings and counts for these three types of omissions.

Number of omissions (reported after pictures).

One piece of evidence that supports existence of the three types of omissions is that the participants indeed chose to make use all three of the options we provided. First, participants reported a number of cases in which they remembered some information during initial recall but simply neglected to write it down at the time (*neglected* = 27). The presence of the *neglected* response option in the computer interface for each timeslice also may have encouraged participants to report as fully as possible on subsequent timeslices since they would realize that they were going to have to type in any omitted information anyway. Second, participants reported a substantial number of cases in which the pictures reminded them of their original experience (*reminded = 203)*. Third, participants demonstrated that they were perfectly willing to report when the pictures showed something that they must have experienced but nevertheless could not remember at all (*forgotten* = 60).

Characteristics of the three types of omissions (reported after pictures).

The patterns of the rating data in Table 2to3 lend additional support for distinguishing these three types of omissions.

Of the three types of omissions, *neglected* showed the pattern of ratings most similar to the initial accurate responses. This makes sense because participants were indicating that they had in fact earlier recalled these aspects of their original experience, but just neglected to write them down during initial recall. The neglected omissions showed higher frequency ratings than the initial accurate responses, likely because the most frequent aspects of experience, those that remain constant across many timeslices, were so common that they seemed too trivial to mention during initial recall (e.g., actions: “sitting at a computer”, environment/weather: “sunny”, clothes & belongings: “watch”; see Brewer & Treyens, 1981, pp. 222-226). The neglected omissions also showed higher reliving ratings than the initial accurate responses, likely due to a combination of: their high frequency (perhaps even occurring in other timeslices), the stimulation from seeing the pictures, and the fact that they had ostensibly already been retrieved earlier. The most common of the visible categories for neglected omissions were: actions (M = 29%), locations (21%), people (19%), and clothes & belongings (16%).

The *reminded* omissions showed reliving ratings similar to the initial accurate responses, along with the lowest knowledge and frequency ratings. The low knowledge ratings make sense, because these are generally uncommon aspects of experience that participants have just been reminded of and are reliving. Thus the reports of these omissions do not come from participants’ general knowledge but from recollection. The low frequency ratings of the remindings reveal a similarity between our task and more standard laboratory recognition tasks. Evidence from laboratory list learning experiments shows that low frequency words are more likely to be recognized than high frequency words (see reviews in Kintsch, 1970 and Mandler, 1980). In a more naturalistic task of incidental memory for a room, Brewer and Treyens (1981) found that unexpected items (low frequency) showed a pattern of stronger recognition than would be expected from their recalls. In a similar way, our reminded omissions were aspects of experience that were not recalled during the initial recall phase, but were nevertheless recognized and relived once participants saw the pictures. The most common of the visible categories for reminded omissions were: actions (M = 38%), locations (17%), visual & spatial (14%), people (14%), and information content (12%).

The reminded omissions were also the longest type of omission report (M = 11.1 words, vs. 4.0 for neglected, and 5.2 for forgotten), even longer than the initial accurate responses (7.2). They featured much more idiosyncratic detail than the other types of omissions. Here are examples of reminded omissions from three different participants: (1) “Kids in a car threw what felt like a pine cone at me that hit my arm as I was crossing Goodwin on Green” (2) “Video that I was showing people in my phone was of a squirrel trying to lick the inside of an empty milkshake container” (3) “Watched as a baby bird flew from its nest and sat on the steps to keep an eye on it”. Compare those to the longest neglected omissions from the same participants: (1) “Paid for the nectarines at an automated cash register” (2) “Sandals” (3) “Dark outside, nice temperatures”. And compare also to the longest forgotten omissions from the same participants: (1) “TV in front of me as I sat on the couch” (2) “I could see the IGB Gatehouse from where I was sitting." (3) "Went on my laptop". The rating profile of high reliving, low knowledge, and low frequency, combined with the response length and detail, illustrate *the power of pictures from a wearable camera to evoke recollection of specific episodes*. In the next section we will furthermore consider remindings of *nonvisual* aspects of experience.

Finally, the *forgotten* omissions were cases in which the participants reported having no memory of an aspect of experience that the pictures indicate must have occurred in their lives. These omissions showed low reliving ratings, high knowledge ratings, and high frequency ratings. The high frequency ratings are consistent with the literature discussed above showing that higher frequency items are *less* likely to be recognized. The high knowledge ratings may reflect the fact that the only reason participants knew these aspects of experience happened was because the pictures showed them, not because they remembered them. The most noteworthy characteristic of the forgotten omissions is their low reliving ratings. There is only limited evidence on the relationship of reliving to memory accuracy. Our own data from initial recalls show a negligible reliving/accuracy relationship (Table 2). However, several other studies have found a positive relationship. Brewer's (1988) beeper experience sampling study of autobiographical memory showed a positive relationship (see Table 3.15, p. 68). Morris (1992) and Robinson et al. (2000) carried out studies of memory for filmed events that showed a positive relationship between mental imagery and recall accuracy. If we conclude that reliving/imagery is positively related to memory accuracy, then items that were reported as forgotten would be expected to show low scores on reliving, and that is indeed what we found. The most common of the visible categories for forgotten omissions were: actions (*M* = 44%), locations (17%), visual & spatial (13%), and clothes & belongings (12%).

There was no difference in certainty ratings across the three types of omissions. This is expected. The certainty ratings were all at ceiling, because the pictures provided objective proof of the original experiences.

Overall, this analysis of the rated characteristics of the omissions, in light of previous memory findings, supports the division of the omissions into the three separate types (*neglected*, *reminded* and *forgotten*) and provides some insights into the different memory processes associated with each type of omission. We think it is important to be able to distinguish between these three qualitatively different types of omission reports evoked by the pictures, and these distinctions proved useful in our calculation of a memory completeness measure. Any future studies that elicit recall both before and after viewing photos or videos should include such a distinction in their procedure.

Kintsch, W. (1970). *Learning, memory, and conceptual processes*. New York, NY: John Wiley & Sons.

Mandler, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review*, *87*(3), 252-271. https://doi.org/10.1037/0033-295X.87.3.252

Considerations in the Calculation of Completeness in Autobiographical Memory

Memory completeness is the proportion of original experience remembered. The challenge with such a measure for autobiographical memory is that we must decide how to quantify the participants' original experiences in order to form the denominator. In a laboratory free recall experiment, this is straightforward: if there were 20 words on a list, the denominator is 20. But what is the totality of everyday experience? What would a 100% complete memory of an ordinary day's activities be like? Would it be the unaided replay of all the activity of every neuron, such that remembering would be like reliving the entire day in every last detail, in real time, as does the character in the fictional work *Funes the Memorious* by Jorge Luis Borges (1942)? In such case, the denominator might as well be infinity.

Can we circumscribe our completeness denominator by considering the maximum amount of incoming information that can be perceived and rapidly encoded? Despite the massive yet difficult-to-quantify capacity of human long-term memory (cf. Brady et al., 2011; Dudai, 1997), research on change blindness argues against too expansive a view of the amount of information that is actually available for autobiographical memory (Simons & Ambinder, 2005). Change blindness research has shown that people sometimes fail to notice substantial changes to a visual scene, such as the person they are talking to being replaced by another person after a brief interruption (Simons & Levin, 1998). Such findings suggest that people make use of only a subset of the information impinging on their senses; that is, the amount of information that actually makes it into our memory representations is likely much less than we are led to believe from the apparent richness of our phenomenal experiences. One might think to use estimates of the functional input rate of human experience to constrain our analysis. However, such estimates range so widely that they are of little help (e.g., 1.2 bits per second, Landauer, 1986; 10^8 bits per second, Koch, 1997). Even if there were a consensus about the amount of incoming information experienced in a particular duration of time, we could still not use it since we do not know how to convert participants' recall of everyday events into quantities of bits without first understanding the representational structure of everyday experiences.

As we see it, our best recourse is to use actual records of the original experience (i.e., external memory; Finley et al., 2018) to provide an estimate of the total amount of information, and our participants themselves are the ones best suited to extract as much information as possible from those records. C. S. Lewis (1967) claimed: "A single second of lived time contains more than can be recorded." Thompson et al. (1996, p. 26) claimed: "any record will abstract and condense the event." We are not in a position to dispute these claims at present. Thus we do not attempt here to define what 100% complete recall of a timeslice of everyday activity would be. Instead, we take a much more conservative and practical approach, so that we may at least calculate an *upper bound estimate* of completeness. Thus, the question we are asking is: how much were participants able to remember, unaided, about their original experiences compared to how much they were able to report about those experiences when provided with a pictorial record? We understand that this approach yields a measure of completeness that is clearly too large, because the total amount that participants can report with the help of the pictures must be considerably smaller than the totality of their original experience. Nevertheless we think it is important to begin providing data on this issue, even if all we can do is suggest an upper bound for the completeness of autobiographical memory for everyday experiences.

The numerator in our formula for completeness consists of the count of initial recall responses (i.e., those made before seeing a timeslice's pictures) plus omissions classified as *neglected* (i.e., those that participants stated they had remembered before seeing the pictures but had simply neglected to report). The denominator is the count of initial recall responses plus *all* types of omissions (*neglected*, *reminded*, and *forgotten*). We then need a method to handle the counting of *erroneous* initial recall responses. We have chosen to dichotomize the initial recall responses into accurate (self-scored 7 on the scale of 1-7) versus inaccurate (self-scored 6 or lower), and only include accurate responses in our formula. By analogy, in a laboratory free recall experiment, one would not include intrusions (i.e., recalled words that were not on the list) when calculating the proportion of the original list that a participant recalled. Thus, our formula for completeness is:

$$completeness = \frac{(accurate\ initial\ recalls + neglected\ omissions)}{(accurate\ initial\ recalls + all\ omissions)}$$

Borges, J. L. (1942, May 7). Funes el memorioso [Funes the Memorious]. *La Nación*. Buenos Aires, Argentina.

Brady, T. F., Konkle, T., & Alvarez, G. A. (2011). A review of visual memory capacity: Beyond individual items and toward structured representations. *Journal of Vision, 11*(5). https://doi.org/10.1167/11.5.4

Dudai, Y. (1997). How big is human memory, or on being just useful enough. *Learning & Memory, 3*, 341-365. https://doi.org/10.1101/lm.3.5.341

Koch, C. (1997). Computation and the single neuron. *Nature 385*, 207-210. https://doi.org/10.1038/385207a0

Landauer, T. K. (1986). How much do people remember? Some estimates of the quantity of learned information in long-term memory. *Cognitive Science 10*, 477-493. https://doi.org/10.1207/s15516709cog1004_4

Lewis, C. S. (1967). Historicism. In W. Hooper (Ed.) *Christian Reflections* (pp. 100-113). Grand Rapids, MI: Eerdman's Publishing Company.

Simons, D. J., & Ambinder, M. S. (2005). Change Blindness: Theory and Consequences. *Current Directions in Psychological Science, 14*(1), 44-48. https://doi.org/10.1111/j.0963-7214.2005.00332.x

Simons, D. J., & Levin, D. T. (1998). Failure to detect changes to people during a real-world interaction. *Psychonomic Bulletin & Review, 5*(4), 644-649. https://doi.org/10.3758/BF03208840

Methods Used in the Comparable Studies of Autobiographical Memory (Figure 4)

Where necessary, values were extracted from published figures using Engauge Digitizer version 11.3 (doi:10.5281/zenodo.3344601).

Linton (1975, 1978, 1982) recorded short verbal descriptions for approximately two *distinctive* events at the end of each day for about six years. When she tested herself on events, monthly, she marked events as forgotten if she could not recall or could not distinguish them from other events in memory. The completeness measure we used for her study was the proportion of tested events not marked as forgotten (Linton, 1982, Figure 11-2).

Wagenaar (1986) followed a procedure similar to Linton's, recording descriptions for approximately one distinctive event at the end of the day for about four years, and marking an event as forgotten at test if he was not reminded of it in particular even upon reading all of the event's recorded details. The completeness measure we used for his study was the proportion of tested events not judged forgotten (Wagenaar, 1986, Figure 3).

White (1982, 1989, 2002) recorded descriptions of one event per day, with events chosen "haphazardly" (sometimes ordinary and sometimes unique), for one year. At test, he looked at an event's complete record and rated his memory for it on a scale of 1 to 5, where 1 indicated no recollection and 5 indicated total recall. The completeness measure we used for his study was the proportion of tested events given ratings higher than 1 (White, 2002, Table 1).

In Brewer's 1988 study, 18 undergraduate participants (Experiments 1 and 2 combined) recorded descriptions of events as prompted *randomly* throughout the day (approximately 3 to 5 times per day) for approximately two weeks. Eight of those participants (Experiment 1) additionally recorded descriptions of the day's one most memorable event at the end of each day. At test, participants were cued by one aspect of an event's record (e.g., actions, thoughts) and they rated their memory for the event on a scale of 1 to 7, where 1 indicated "have no memory of the event" and 7 indicated "certain that remember event". In the data reported, Brewer chose to classify responses of 1 or 2 as forgotten. The completeness measure we used for this study was the mean proportion of tested events given memory ratings of 3 or higher when cued with the event's action information, which was the most effective cue (Brewer, 1988, Figures 3.2, 3.5). This measure is likely slightly inflated compared to measures from some of the other studies, due to the original classification treating responses of both 1 and 2 as forgotten, rather than just 1. The original data were not easily available for reanalysis, so we used the data as reported in the publication.

Thompson and colleagues (1996) conducted a large number of daily diary studies with undergraduate students as participants. The studies were chiefly concerned with participants' estimations of the dates of events, but also included memory measures suitable for comparison to the other studies included here. In most of the studies by Thompson et al., participants wrote a verbal description of one event per day for 3.5 months (and up to 2.5 years in one study). Participants were instructed to choose a *variety* of events, some memorable and some not memorable, some pleasant and some unpleasant. They were also instructed to "strive to identify the events uniquely in the diary records" (p. 24). Non-unique recorded events were excluded from testing (p. 36). At the end of acquisition, participants were cued with their event descriptions in a random order. For each event they made a memory rating on a scale of 1 to 7, where 1 indicated that they did not remember the event at all and 7 indicated that they remembered the event perfectly. In the long duration study, participants were also asked to recall location and people when cued with the entry. The completeness measure we used for the ratings was the proportion of events given ratings of 2 or higher, collapsed across participants (Thompson et al., 1996, Table 3.5, $N = 428$, retention intervals of 1, 5, and 10 weeks). For the recall of location, we used percent correctly recalled (Figure 1.1, $N = 6$, retention intervals from 100-700 days). Presumably the data points in that original figure were the means across participants. Recall of people was similar, and we did not include it in our figure.

In the study by Finley, Brewer, and Benjamin (2011), 12 undergraduate participants wore SenseCam for five days. On a given day, the camera took pictures as triggered by the timer (approximately every 10 s) or as triggered by the sensors. Trigger condition had no statistically significant effect, so we collapsed across it for use of the data in Figure 4. Furthermore, on a given day participants either reviewed their pictures at the end of the day or had no such review. For Figure 4, we only used data from the no-review days, in order to maintain similarity to the current study. In the 2011 study participants were tested after retention intervals of 1, 3, and 8 weeks. There were several dependent measures used on the tests. The one we focused on here was a recognition rating: participants were shown one of their pictures and rated the strength of their memory for the scene and/or event depicted by the picture on a scale of 1 to 7, where 1 indicated “no memory” and 7 indicated “extremely strong memory.” The completeness measure we used for this study was the proportion of pictures given ratings of 2 or higher. We calculated this for each participant from the original raw data, and plotted the mean.

For the current study, we used completeness as calculated using Formula 1, for the visible response categories only, separately for test 1 (one week retention interval) and test 2 (one month retention interval). We calculated completeness for each participant and plotted the mean. Note that this measure did not significantly differ from test 1 ($M = .81$, $SD = .12$) to test 2 ($M = .75$, $SD = .12$), $t(8) = 1.50$, $p = .173$, $d = 0.43$.

Reanalysis of Brewer (1988) Autobiographical Memory Error Rates, and Comparison to Current Study

We re-analyzed data from Brewer (1988, Table 3.8) in order to compare the error rate for autobiographical memory from that study to the error rate in the current study. In this re-analysis we selected the subset of data from the earlier study that most closely corresponded to the data in the current study. We used only data from Brewer's Experiment 2 (which used recall, whereas Experiment 1 used recognition), and only data from the recall of *Actions* since that was consistent with our decision in the current study to examine only potentially visible categories. As shown in Table 3.8, there were 533 Action event recalls in Brewer's Experiment 2, combining across three tests with mean retention intervals of 7, 30, and 53 days. In that experiment, the event recalls were classified by the experimenter into 7 recall categories after comparing the initial description of the event with the recall of the event (Correct with Detail [22]; Correct [109]; Inference [70]; Overt Error [4]; Wrong Time Slice [20]; Wrong Event [118]; Omit [190]). In the current study we excluded responses for "miscued timeslices" (i.e., when participants reported that the verbal timeslice cue led them to recall a totally different timeslice from that shown in the pictures). Therefore in our re-analysis of the data from Brewer (1988) we excluded recalls classified as Wrong Event. Omits were excluded from the error analysis of both studies. In the current study, error responses were those self-scored as less than totally accurate (<7 on the 1-7 accuracy scale) by the participants after comparing their recall with the evidence in the relevant pictures; these error responses were further sorted into classes by the two authors of the current study. The two recall categories from Brewer (1988) that correspond to the ones in the current study were Overt Errors [4] and Wrong Time Slice [20]. Thus in the re-analysis of Brewer (1988), for the recall data closest to the current study, there were a total of 24 recall errors out of 225 event recalls giving an error rate of 10.67%.

Finally, aspects of a study's methodology may influence the chances of eliciting or detecting certain types of errors. Comparing errors in the current study to Brewer's (1988) beeper study, we find that the current study had a higher proportion of *substitutions* and *distortions*, while the beeper study had a higher proportion of *time shift* errors (referred to as "wrong time slice"). In the beeper study's testing sessions, participants were cued with one part of the response cards they had recorded during acquisition: time, location, time and location, thought, or action. So a majority of the cue types gave no temporal information, leaving much room for time shift errors. In contrast, in the current study's testing sessions, participants were cued with the date, time, and verbal description of the start and end of a timeslice. This cueing procedure probably reduced the number of time shift errors, though they still were the most frequent error category in our data. In the beeper study, any *distortions* or *substitutions* were likely harder to detect because participants' responses were narrative descriptions, both at acquisition and test, and they may not have thought to write about the same aspects of experience in enough detail for an error to be noticed (e.g., mentioning the shape of a table both at acquisition and test). In contrast, the procedure of the current study encouraged detailed responses across the entire range of aspects of experience, and initial recall responses were compared to pictures for scoring, thus making it easier to notice distortions or substitutions. Different methodologies will each have their own strengths and weaknesses.

Commentary on Retention Interval

Our primary reason for including two retention intervals in this study (approximately 1 week and 1 month) was to safeguard against possible floor or ceiling effects in the initial recall data. Examination of the data showed no obvious floor or ceiling effects so we collapsed the data across the two retention intervals in all subsequent analyses. That is, for each participant we combined data points from both testing sessions and proceeded with analysis as if there had been only one testing session.

Data reported separately for the two tests can be found below in Tables S1 and S2. Retention interval (test 1 vs. 2) did not have a statistically significant effect on any of our main dependent measures. Accuracy (proportion): $t(8) = 0.20$, $p = .850$, $d = 0.10$. Accuracy (rating): $t(8) = 0.49$, $p = .636$, $d = 0.26$. Completeness: $t(8) = 1.50$, $p = .173$, $d = 0.42$. None of the initial recall ratings (certainty, reliving, knowledge, and frequency) differed significantly across the two tests. Thus we are not losing meaningful information by collapsing across retention intervals. Basic characteristics of responding that did differ across testing sessions included the number of responses (test 1 > test 2) and the number of timeslices completed (test 1 < test 2), as shown in Table S1. Those results show that there was some forgetting across the three weeks between the tests. Why would this not be reflected in our accuracy and completeness measures? Participants' self-scoring of their initial recall responses, and reporting of omissions, were based on the information available directly in the pictures as well as participants' own memories as further stimulated by the pictures. Thus, their ability to self-score and report omissions necessarily declined over time along with their ability to retrieve information from memory for initial recall, and likely at the same rate. There is also the simple issue of power. Our measure of completeness did numerically decrease from test 1 (*M* = .81, *SD* = .12) to test 2 (*M* = .75, *SD* = .12), but not enough for statistical significance with our modest sample size.

Table S1

Timeslice Data by Test Number

	Test 1	Test 2	Test 1 vs. 2		Overall
Measure	*M (SD)*	*M (SD)*	*t*	*p*	*M (SD)*
Timeslices created	12.0 (3.5)	10.8 (3.7)	1.42	.194	11.4 (3.4)
Proportion timeslices completed on test	.66 (.26)	.85 (.23)	2.38	.045	.76 (.21)
Proportion miscued timeslices	.12 (.19)	.14 (.22)	0.64	.541	.13 (.20)
Time covered per test (hours)	5.2 (2.1)	6.8 (3.6)	1.01	.344	6.0 (2.0)
Initial recall responses per timeslice	15.3 (8.0)	10.3 (5.8)	3.44	.009	12.5 (6.3)
Omissions per timeslice	4.0 (3.4)	3.2 (2.9)	1.83	.105	3.5 (3.1)
Total responses per hour original experience	48.2 (26.2)	29.4 (27.1)	2.31	.050	37.7 (24.1)
Initial recalls word count	8.4 (5.2)	7.2 (4.8)	2.09	.070	7.7 (5.2)
Omissions word count	25.7 (11.8)	22.8 (17.9)	0.12	.910	36.6 (22.1)
Timeslice overall memory strength rating (1-7) before pictures	4.1 (0.9)	3.5 (1.2)	2.17	.062	4.12 (0.74)
Timeslice overall reliving rating (1-7) before pictures	3.6 (1.0)	3.0 (0.9)	1.36	.211	3.58 (0.66)
Timeslice overall memory strength rating (1-7) after pictures	4.9 (0.9)	4.6 (1.1)	0.99	.350	4.88 (0.71)
Timeslice overall reliving rating (1-7) after pictures	4.3 (1.2)	4.2 (1.0)	0.31	.764	4.40 (0.87)

Note. Retention interval was 1 week for Test 1 and 1 month for Test 2. All rows below “Proportion miscued timeslices” were calculated excluding miscued timeslices. Omissions were reported by participants after viewing a timeslice’s pictures.

Table S2

Means (and Standard Deviations) of Performance Measures Across Tests 1 and 2

	Test 1	Test 2	Test 1 vs. 2		Overall
Measure	*M (SD)*	*M (SD)*	*t*	*p*	*M (SD)*
Completeness	.81 (.12)	.75 (.12)	1.50	.173	.79 (.11)
Memory Increase Due to Pictures	.19 (.12)	.28 (.19)	1.57	.154	.23 (.14)
Information Increase Due to Pictures	.26 (.19)	.36 (.20)	1.72	.123	.29 (.18)
Accuracy (Proportion)	.88 (.11)	.89 (.06)	0.20	.850	.89 (.06)
Accuracy (Rating)	6.46 (0.48)	6.57 (0.31)	0.49	.636	6.52 (0.26)
Response Ratings					
Certainty	6.03 (0.65)	6.15 (0.54)	1.15	.285	6.10 (0.58)
Reliving	4.29 (1.04)	4.00 (1.42)	0.60	.566	4.21 (1.00)
Knowledge	4.15 (1.22)	4.19 (1.48)	0.12	.908	4.16 (1.27)
Frequency	4.37 (0.86)	4.47 (1.02)	0.36	.731	4.41 (0.83)

Note. Retention interval was 1 week for Test 1 and 1 month for Test 2. Means of participant means, for potentially visible response categories only. Rating measures were all made on 1 to 7 scales.

Table S3

Means (and Standard Deviations) of Ratings for Initial Responses by Response Category

Response Category	Response Count	*n*	Ratings at Time of Recall (Before Pictures) Certainty	Reliving	Knowledge	Frequency	Accuracy Rating	Words per Response
Potentially Visible								
Actions	199	9	6.0 (0.6)	4.3 (1.0)	3.9 (1.2)	4.1 (0.9)	6.1 (0.7)	9.4 (7.1)
Clothes & Belongings	184	9	6.2 (0.5)	3.9 (1.2)	4.5 (1.5)	4.7 (1.0)	6.1 (0.8)	5.7 (3.3)
Environment/Weather	109	9	5.7 (0.7)	3.8 (1.2)	4.3 (1.2)	5.0 (1.1)	6.5 (0.2)	4.7 (1.8)
Information Content	53	7	6.1 (0.9)	4.8 (0.7)	3.2 (1.4)	2.6 (0.6)	6.3 (0.9)	16.3 (7.2)
Locations	212	9	6.3 (0.5)	4.2 (1.0)	4.2 (1.3)	4.3 (1.0)	6.6 (0.4)	6.2 (3.5)
People	185	9	6.1 (1.0)	4.4 (1.2)	4.0 (1.3)	4.4 (0.9)	6.7 (0.4)	4.5 (2.9)
Visual & Spatial	143	9	6.1 (0.8)	4.6 (0.9)	4.2 (1.3)	4.5 (0.9)	6.0 (0.9)	12.5 (7.8)
Not Potentially Visible								
Audio	100	8	5.6 (1.2)	4.2 (1.4)	4.0 (1.6)	4.9 (1.1)	6.4 (0.5)	7.2 (6.3)
Feelings	106	9	5.9 (1.0)	4.6 (1.0)	3.6 (1.3)	4.0 (1.3)	6.8 (0.2)	8.4 (6.6)
Thoughts	87	8	6.0 (0.8)	4.8 (1.0)	3.3 (1.5)	3.5 (1.1)	6.6 (0.5)	13.1 (7.5)
Touch, Taste, & Smell	82	9	5.0 (1.6)	3.4 (1.3)	3.8 (1.2)	4.4 (1.1)	6.3 (0.4)	8.5 (6.8)

Note. Rating data are means of participant means, collapsed across retention intervals, and excluding miscued timeslices; *n* is the number of participants (out of 9) who gave each type of response; rating measures were all made on 1 to 7 scales.

Table S4

Repeated Measures Correlations Among Initial Recall Response Ratings with 95% Confidence Intervals

Response Rating Measure	Confidence	Reliving	Knowledge	Frequency	Accuracy
Confidence	--				
Reliving	**.35 [.29, .40]**	--			
Knowledge	**-.07 [-.13, -.01]**	**-.55 [-.59, -.51]**	--		
Frequency	.01 [-.05, .07]	**-.31 [-.36, -.25]**	**.36 [.30, .41]**	--	
Accuracy	**.26 [.20, .31]**	*.05 [-.01, .11]*	.01 [-.05, .07]	**.06 [.00, .12]**	--

Note. Data are only from potentially visible response categories, collapsed across retention intervals; rating measures were all made on 1 to 7 scales; accuracy was self-scored after seeing pictures; $N = 9$; $df = 1{,}075$; **boldface**: $p < .05$; *italics*: $p < .10$; 95% confidence intervals in brackets; see Bakdash & Marusich (2017) for specification of repeated measures correlation.

Table S5

Mean Pearson Correlations Among Initial Recall Response Ratings

	Certainty	Reliving	Knowledge	Frequency	Accuracy
Certainty	--				
Reliving	**.39 (.10)**	--			
Knowledge	.04 (.32)	*-.32 (.43)*	--		
Frequency	.04 (.13)	-.21 (.33)	**.41 (.15)**	--	
Accuracy	**.27 (.14)**	*.07 (.09)*	.03 (.13)	.03 (.16)	--

Note. $N = 9$; Pearson correlations were calculated separately for each participant using only data from potentially visible response categories, collapsed across retention intervals; means displayed in table; standard deviations are in parentheses; data are only from potentially visible response categories, collapsed across retention intervals; rating measures were all made on 1 to 7 scales; accuracy was self-scored after seeing pictures; **boldface**: $p < .05$; *italics*: $p < .10$; Table 2 in the main paper reports repeated measures correlations.

In the main paper, we reported repeated measures correlations which account for individual differences without sacrificing power (Bakdash & Marusich, 2017). Here we show that the same pattern of results holds if we instead calculate separate Pearson correlations for each participant and then calculate the average.

Additional Data on Characteristics of Episodic and Semantic Autobiographical Memory Response Categories

To what extent were responses in different categories more episodic or semantic? Here we include even the nonvisible categories, as we are not concerned with accuracy ratings. For each response category, we calculated an episodic/semantic score (reliving minus knowledge), collapsing across testing sessions and averaging across participants (see also Table S2). Positive values are more episodic and negative values are more semantic. In descending order from most episodic to most semantic, the results were: Information Content (1.53); Thoughts (1.52); Feelings (0.98); People (0.39); Visual & Spatial (0.36); Actions (0.35); Audio (0.13); Locations (-0.02); Touch, Taste, & Smell (-0.42); Environment/Weather (-0.47); Clothes & Belongings (-0.61). Note that there appears to be a cluster of categories on both ends of the scale. The three most episodic categories (information content, thoughts, and feelings) seem to be aspects of everyday experience that are most likely to fluctuate across episodes. Information content and thoughts were also the two categories with the longest mean response lengths (16.3 and 12.1 words per response, respectively; see Table S2). These aspects of experience were distinct and recalled with detail, thus supporting their characterization as episodic. Here we provide a few examples, each from a different participant, that furthermore illustrate the richness of some of the recalls from the nonvisible categories. In examples in this section, we use the abbreviations *R* for reliving rating, *K* for knowledge rating, and *F* for frequency rating.

> Information Content: “We talked about dinner yesterday. [R] didn't know what Red Lobster was so [N] explained it to her and then [N] commented on how the food was a little too salty. But I said that's how most Americans prefer pasta and she agreed, using a story about a friend from high school as an example.” ($R = 7$, $K = 1$, $F = 1$)

Thoughts: “I was thinking about how I was feeling uncomfortable to be sitting on the grass because there were ants crawling around me.” (*R* = 6, *K* = 1, *F* = 4)

Feelings: “I was nervous that something might go wrong with activating or acquiring our phones. I was excitied from finally getting my hands on the iPhone. I was happy that the people there were so helpful.” (*R* = 6, *K* = 2, *F* = 1)

Two of the most semantic response categories (environment/weather, clothes & belongings) seem to be aspects of everyday experience that are least likely to fluctuate across episodes (e.g., “Keys, phone, wallet” *R* = 1, *K* = 6, *F* = 7). Thus it makes sense that recall of them would be more likely based on general knowledge. Some information could even be inferred directly from the time of day given in the verbal cue (e.g., “Starting to get dark” *R* = 1, *K* = 7, *F* = 6, for a timeslice around 7:50pm).

Touch/taste/smell responses were in fact almost evenly distributed between episodic and semantic (*n* = 38 vs. 37), but the semantic ones were slightly more semantic than the episodic ones were episodic. Episodic touch/taste/smell responses were often linked to the action or location. For example, one participant had friends over to his place on a Friday night to play a board game, Settlers of Catan, and his attempt to cook a frozen pizza went awry when he left the pizza cooker unplugged (feelings included “Embrassed”). The touch/taste/smell response in this timeslice was: “Pizza tasted ok but smelled a little burnt on top.” (*R* = 4, *K* = 2, *F* = 4). Compare this to a semantic touch/taste/smell response, many of which were based on schemas or general knowledge: “Breakfast was good as usual. A little sweet if I had cereal because I usually add some honey and cinnamon sugar.” (*R* = 2, *K* = 6, *F* = 7).

Response Length

Back to considering only the visible categories, the length of episodic responses (*M* = 8.1 words, *SD* = 5.1) tended to be greater than the length of semantic responses (*M* = 6.0, *SD* = 3.0), $t(8) = 2.16$, $p = .063$, $d = 0.48$. Furthermore, there was a positive correlation between response length and episodic/semantic score (reliving minus knowledge), $r_{rm}(1075) = .17$, $p < .001$. These patterns remained the same if we included the nonvisible categories in analysis. Participants simply had more to say about their episodic responses versus their semantic ones.

Other Characteristics

Table S5 shows the means (averaged across testing sessions, then across participants) and standard deviations of several measures for episodic versus semantic responses, for the potentially visible response categories only. Note that accuracy was high for both classes, and did not reliably differ across classes. This suggests that when participants are queried about their own everyday experiences, there are two qualitatively different ways that they can generate accurate responses.

Table S6

Statistics for Episodic Versus Semantic Responses

Measure	Episodic	Semantic	Episodic vs. Semantic	
	M (SD)	*M (SD)*	*t*	*p*
Count Per Test	32 (28)	23 (15)	0.77	.462
Accuracy (binary)	.80 (.18)	.76 (.15)	0.82	.435
Accuracy Rating (1-7)	6.4 (0.5)	6.3 (0.3)	0.89	.402
Certainty About Accuracy	6.7 (0.4)	6.5 (0.4)	2.80	**.023**
Initial Response Ratings				
Certainty	6.4 (0.6)	5.7 (0.4)	3.68	**.006**
Frequency	3.6 (0.8)	4.9 (0.9)	3.89	**.005**
Correlations				
Certainty x Accuracy	.04 (.20)	.48 (.24)	2.84	**.025**
Certainty x Frequency	.05 (.13)	.11 (.15)	0.71	.498
Frequency x Accuracy	.10 (.15)	.07 (.20)	0.25	.812

Note. We classified a response as episodic if its reliving rating was higher than its knowledge rating, and we classified a response as semantic if its knowledge rating was higher than its reliving rating

UNIVERSITY OF ILLINOIS
AT URBANA-CHAMPAIGN

Department of Psychology
College of Liberal Arts and Sciences
603 East Daniel Street
Champaign, IL 61820

Informed Consent

Please read this consent agreement carefully. You must be 18 years old or older to participate.

This research study is being conducted by faculty of the UIUC Psychology Department (Professors William F. Brewer and Aaron S. Benjamin) and a graduate student of the UIUC Psychology Department (Jason R. Finley).

Purpose of the research:

The purpose of this research is to develop a better understanding of the functioning of human autobiographical memory and to use this new knowledge to improve the lives of individuals with disorders of memory. This research will make use of SenseCam, a small wearable low-resolution digital camera that can passively take pictures throughout the day, at fixed intervals, or in response to changes in light levels, or when manually triggered by the individual wearing the SenseCam. SenseCam has been developed by the UK Research Unit of Microsoft, which has provided us with several SenseCams for our memory research.

The goal of our study is to use this new technology to study the characteristics of autobiographical memory (e.g., which types of events are best recalled) and to examine the impact of a daily review of the events captured by SenseCam on overall memory for events in one's experience that day. These results can then inform the work of the Microsoft, UK Research Unit that is studying the use of SenseCam as a therapeutic aid for patients with memory disorders.

What you will do in this study:

You will wear a SenseCam as you go about your normal daily activities, for 1 to 5 consecutive days. On the evening before the first day, you will come to the laboratory (in the UIUC Psychology Building) and meet with me to receive and become familiarized with the SenseCam. You will also be asked to provide information about their typical weekly schedule. You will take the SenseCam home with you, turned off, and will turn it on and wear it beginning the next morning. The SenseCam will either operate automatically throughout the day, or you will be instructed to periodically press the shutter button in order to manually trigger the SenseCam. You may also be asked to periodically write down information about the present event (e.g., location, time, thoughts, feelings, actions). After wearing the SenseCam that first day, or after several days, you will return to the laboratory in the evening and meet me and I will download the images from the SenseCam onto an external hard drive, and delete, without seeing, any images falling in a time period that you have marked for deletion. You will then either watch a review of images taken throughout that day (in the form of a simple "movie," composed of a sequence of images), or you will receive no such review. You may also be asked to mentally review the events of the day (that is, to simply think over the day's events in your head), write down the day's events, and/or make judgments about some of the day's images (e.g., rating them for memorability). You will then take the camera home with you, turned off, and plug it in to be recharged overnight. This daily procedure will be repeated for a total of five days. You will return the SenseCam to me on the evening of the final day. You will then be scheduled to return to the laboratory for 1 to 4 subsequent testing sessions, at intervals of days, weeks, and/or months.

During each testing session, your memory for events of the SenseCam days will be tested in one or more of the following ways:

- You may be asked to write down details about the events that occurred during the days on which you wore the SenseCam, either cued by images from those days or with no such cues. You may also be cued by a certain day and time, by textual descriptions of images, or by images that have been modified in some way (e.g., flipped, blurred, partially obscured, or combined with parts of other images).
- You may be asked to judge whether or not you saw the particular event depicted by an image or images.
- You may be asked to evaluate the memorability of events that transpired during SenseCam days.
- You may be asked to rate how vividly you recollect the events depicted by the images.
- You may be asked to arrange a number of images in the temporal order of the events they depict.

A single testing session will last from 1 to 2 hours.

You may also be scheduled to return to the lab for subsequent review sessions (e.g., to watch a movie of your images after a delay of days, weeks, and/or months).

You will also be asked for any general comments about your experience using the SenseCam and taking the memory tests.

Risks and Benefits:

There are no anticipated physical risks, beyond those encountered in daily life, associated with participating in this study. However, you may potentially feel uncomfortable as a result of inadvertently wearing the SenseCam in a situation where it could take a picture that you would prefer not be seen by the experimenters. However, the experimental procedure allows the graduate student researcher to, at your request, delete any such pictures (specified by time) without viewing them, so this risk is minimized.

The following liability statement is included in accordance with UIUC Policy: The University of Illinois does not provide medical or hospitalization insurance coverage for participants in this research study nor will the University of Illinois provide compensation for any injury sustained as a result of participation in this research study, except as required by law.

The benefits to you, from participating in this study, are minimal. You will have the experience of seeing parts of your everyday life a second time, and may gain insight into your memory for everyday events. This is an experience that previous participants in studies in the UK have found enjoyable.

The benefits for society are large. Microsoft, UK is carrying out an extensive program of research using the SenseCam as a therapeutic tool in patients with memory disorders (e.g., Berry, E., Kapur, N., Williams, L., Hodges, S., Watson, P., Smyth, G., Srinivasan, J., Smith, R., Wilson, B., & Wood, K. (2006). *The use of a Wearable*

Camera, SenseCam, as a Pictorial Diary to Improve Autobiographical Memory in a Patient with Limbic Encephalitis. Neuropsychological Rehabilitation).

Our research project is designed to provide the basic research on autobiographical memory using SenseCam that will support the clinical work on therapy for memory disorders being carried out at Microsoft, UK.

Compensation:

You will be paid as follows:

$5 for each day of wearing SenseCam, to be paid in lump sum at the end of the last SenseCam day.

$15 upon completion of the final testing session.

$10 upon completion of each intermediate testing session, if any.

If you choose to withdraw from the experiment, you will receive a prorated payment for the amount you would have expected to receive for the particular period when you withdrew. For example, if there are 5 SenseCam days and you withdraw after completing two of the five SenseCam days, you will receive $10 (two fifths of the $25 you would have received for completing all five days).

Voluntary Withdrawal:

Your participation in this study is completely voluntary, and you may decide not to participate or may withdraw from participation at any time during the study without penalty or loss of benefits to which you are otherwise entitled. You may skip over any questions or procedures, or you may withdraw by informing the research associate that you no longer wish to participate (no questions will be asked). Your decision to participate, decline, or withdraw from participation will have no effect on your grades at, status at, or future relations with the University of Illinois.

Confidentiality & Privacy:

Your participation in this study will remain completely confidential, and your identity will not be stored with your data. From the very beginning of the experiment, you will be assigned an arbitrary code number, and all data (including images, and responses to the memory tests) will be stored in terms of this code number and will not be linked to your name or other identifying information. All data, consent forms, and initial questionnaires will be stored in a locked filing cabinet in a locked room. Digitally collected data (including images) will be stored on an external hard drive that will be stored in a locked filing cabinet in a locked room. The data key (which links the names of participants to their code numbers) will be stored in a locked location, separate from the consent forms and initial questionnaires, that only the experimenters will have the ability to access.

While no participants' images or personally identifying information will ever be publicly presented or published, the *overall* results of this study may be presented at conferences and/or published in books, journals, and/or in the popular media, and/or presented to Microsoft as a research report.

Several aspects of the procedure are designed to safeguard your rights and privacy as a participant. The SenseCam is equipped with two methods by which you will be able to deactivate it at any time, should you ever prefer to avoid the possibility of images being captured in any places, times, or situations.

First, the SenseCam has an ON/OFF button for complete deactivation (off) or reactivation (on).

Second, the SenseCam has a "DO NOT DISTURB" button which will put the camera into "DO NOT DISTURB" mode. In this mode, the camera will remain on but absolutely no images will be captured. The SenseCam will remain deactivated for 7 minutes, and will alert you with a beep 15 seconds before it reactivates. You may press the "DO NOT DISTURB" button again to reset the "DO NOT DISTURB" mode to last for another 7 minutes. You may also reactivate the SenseCam from "DO NOT DISTURB" mode at any time by pressing the manual shutter button.

Finally, should you want any images captured during a certain time period to be deleted without being seen by anyone (e.g., during a time when you would have preferred for the SenseCam to be deactivated but forgot to deactivate it), you will be able to note the time period in a small notepad which we will provide to you, and as soon as you return to the lab, the graduate student researcher will delete—without seeing—all images from that period.

In the event that any of the non-deleted images should contain information concerning any illegal activity, we would adhere to section 4.05 of the 2002 APA Ethical Principles of Psychologists and Code Of Conduct, which states: "(b) Psychologists disclose confidential information without the consent of the individual only as mandated by law, or where permitted by law for a valid purpose such as to … (3) protect the client/patient, psychologist, or others from harm."

Furthermore, several aspects of the procedure are designed to ensure the rights and privacy of other people who may incidentally appear in the low-resolution images captured by SenseCam.

In addition to the SenseCam, you will also carry with you:

1. A reference card, containing:
   a. A prepared statement to read to anyone with questions or concerns about the SenseCam: "I am participating in an experiment on everyday memory. This is a digital camera that automatically captures low-resolution still images throughout the day, which will later be used to test my memory. It does not record audio or full-motion video. Any images captured will not be made public in any fashion and will only be seen by myself, during the memory tests, and by the experimenters. If you would prefer, I can turn off or temporarily deactivate the camera, and/or make a note and have the images just taken deleted without anyone seeing them. I can also provide contact information for the experimenters."
   b. A list of places and situations in which to deactivate the SenseCam
      - Places/Situations:
        - Any restroom
        - Any changing room, locker room, etc.
        - Doctor's office
        - ATM or bank
        - Wherever/Whenever you would prefer for images not to be captured
        - Wherever/Whenever anyone requests deactivating
      - When entering any private residence (including your own residence if you have roommates) OR when entering your workplace, you must first deactivate the SenseCam, and then read the above general statement to the resident(s) or to your supervisor, with the following addition: "The camera is currently deactivated, and I will only activate it with your permission." At your

discretion, you may then ask: "May I have your permission to activate the camera?" Activate the camera if given permission.

- A reminder to reactivate the SenseCam when leaving places and situations in which it was deactivated.

c. contact information for the investigators (email, address, and phone numbers)

2. A small notepad in which to note periods of time for images that you want deleted.

Further information:
If you have questions about this study, please contact Jason R. Finley, graduate student, Department of Psychology, University of Illinois, Champaign, IL 61820. Email: [redacted]; cell phone: [redacted]
Or contact the responsible project investigator: Professor William F. Brewer, Department of Psychology, University of Illinois, Champaign, IL 61820. Email: [redacted]; phone: [redacted].

Who to contact about your rights in this study:
If you have any questions about your rights as a participant in this study, please contact the University of Illinois Institutional Review Board at [redacted](collect calls accepted if you identify yourself as a research participant) or via email at [redacted]

Agreement:
The purpose and nature of this research have been sufficiently explained to me. I have read and understand the above consent form and voluntarily agree to participate in this study. I understand that I am free to withdraw at any time without incurring any penalty. I understand that I will receive a copy of this form to take with me. I am 18 years of age or older, and a UIUC undergraduate.

Signature: ______________________________ Date: __________________

Name (print): ____________________________

UNIVERSITY OF ILLINOIS
APPROVED CONSENT
VALID UNTIL

APR 20 2011

Instructions for SenseCam Participants, Summer 2010

Overview of Your Schedule for this Study

1. Sunday
   a. Meet Jason in room 275, Psych Building.
   b. Receive SenseCam (turned off), charger, and notepad.
   c. Charge SenseCam overnight.
2. Monday
   a. Turn on and wear SenseCam when ready to start your day.
   b. Wear SenseCam throughout the day.
   c. Charge SenseCam overnight.
3. Tuesday
   a. Turn on and wear SenseCam when ready to start your day.
   b. Wear SenseCam throughout the day.
   c. Charge SenseCam overnight.
4. Wednesday
   a. **Meet Jason in room 275 Psych Building at time: ____________________**
   b. Return SenseCam, charger, and notepad.
   c. Schedule dates for the 2 testing sessions.

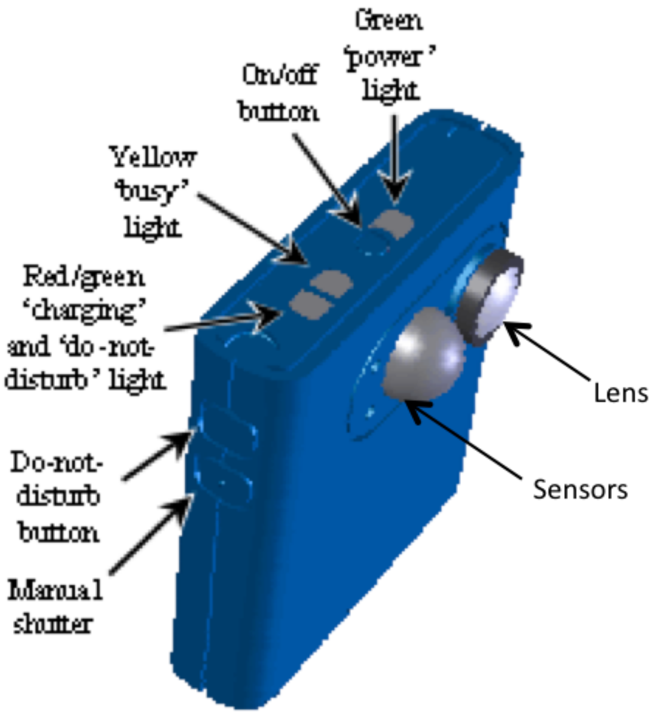
Payment

- $10 cash upon returning SenseCam
- $10 cash at first testing session
- $15 cash at second (final) testing session

Safety while traveling to the Psych Building at night

- Campus Police
  - 217-333-1216
- SafeWalks
  - 217-333-1216
  - The University of Illinois Police Student Patrol offers SafeWalks between the hours of 9:00pm to 1:30am
- SafeRides
  - 217-256-7433
  - http://www.cumtd.com/ridingmtd/services/SafeRide.aspx

<u>Wearing & Operating the SenseCam</u>



Buttons:

- On/off
- Do-not-disturb

Lights:

- Green POWER light:
  - Lit solidly when the SenseCam is on.
  - BLINKING green means LOW BATTERY! Recharge it soon.
- Yellow light:
  - Will blink occasionally while SenseCam i Sensors es NOT correspond to when pictures are being taken.
- Red/Green light:
  - Solid red when SenseCam is in Do-not-disturb mode.
  - Blinks red when about to exit Do-not-disturb mode.
  - Blinking green when SenseCam is charging.
  - Solid green when SenseCam is done charging.

Charging:

1. Leave the SenseCam in its bag while charging, to protect it from getting knocked around. You can plug in the charger through the small hole in the side of the bag, and can reach the ON/OFF button by opening the top of the bag.
2. Press and hold the ON/OFF button until you hear the rising tone. The SenseCam is now booting up. Wait for it to beep once to indicate it has booted.
3. Plug the 2-prong end of the charger into a wall outlet. Make sure the charger's red light turns on.
4. Plug the other end of the charger into the side of the SenseCam. You should hear a falling tone, and you should see a blinking green light on the top of the SenseCam that indicates it is charging. A steady (not blinking) green light means the SenseCam is fully charged.
5. It is fine to leave the SenseCam plugged in overnight.

Starting the SenseCam in the morning (when you are ready to start your day):

1. If you've left the SenseCam plugged in overnight to charge, you should see a solid green light on the top of it, indicating full charge.
2. Unplug the charger from the SenseCam, then from the wall outlet.
3. Press and hold the ON/OFF button until you hear the rising tone. The SenseCam is now booting up. Wait for it to beep once to indicate it has booted.
4. Remove the SenseCam from its bag and put it on by placing the lanyard over your head and around your neck, and adjusting the length of the lanyard to position the SenseCam to high chest level.

Wearing & caring for the SenseCam:

- **Optimal position**: is high chest level (i.e., not down too low, and not so high it's right up against your neck choking you!).
- It should be worn **outside all clothing** (e.g., coat/jacket).
- Be careful to avoid letting any **necklaces, headphone cords, scarves**, **hair,** etc. get in front of the SenseCam or rough it up.
- It's okay to wear while **eating**; just exercise a mild amount of caution to avoid getting food/drink on it.
- SenseCam is **NOT WATERPROOF**, so please be careful to avoid letting it get wet. If it's raining or snowing heavily, it's probably best to tuck the SenseCam inside your coat/jacket while outside (or put it away into its bag until you get our of the rain/snow). If there's just a mild rain or snow and you've got an umbrella, then it should be fine to leave the SenseCam on and outside your coat/jacket.
- Be careful when putting on a **seatbelt** or any kind of over-the-opposite-shoulder bag. Make sure the SenseCam goes over (on top of) the seatbelt or strap, rather than getting smashed.
- Mild to moderate **physical activities** should be pretty safe for the SenseCam (remembering to consider the weather and the not-waterproof part!). Just use your good judgment in considering the safety of the SenseCam. Examples: jogging or bicycling should be fine; racquetball or fencing would be bad!

Things to carry with you each day:

1. SenseCam
2. Notepad
3. SenseCam bag
4. (Also, remember to bring the charger back to the lab when you return the SenseCam.)

Operating the SenseCam (& Notepad):

1. **ON/OFF Button: top of SenseCam**
    a. **Turn ON:** Press and hold the ON/OFF button until you hear the rising tone. The SenseCam is now booting up. Wait for it to beep once to indicate it has booted.
    b. **Turn OFF:** Press and hold the ON/OFF button until you hear the falling tone. The SenseCam is now off.
    c. *Note: If you're going to have the SenseCam off for a while, it may be better to put it in its bag for protection. But please take care not to leave it someplace where it could be stolen (which is pretty much anywhere).*
2. **DO NOT DISTURB Button: upper side button**
    a. Press once to activate "Do Not Disturb" mode (DND). In this mode, the SenseCam will remain on, but absolutely no images will be captured. The SenseCam will remain in this mode for 7 minutes, and will alert you with a beep 15 seconds before it reactivates. You may press the DND button again to renew DND mode to last for another 7 minutes. You may renew DND mode as many times as you want.
    b. You may also end DND mode at any time by pressing the manual shutter button (lower side button).
    c. *Note: DND mode is ideal for using the restroom.*
3. **Manual Shutter Button: lower side button**
    a. This button is normally used to manually trigger the SenseCam to take a picture, but it has been disabled for this study.
    b. The only purpose this button serves in this study is to manually end DND mode.

4. Notepad
    a. If you realize too late that you would've like to deactivate the SenseCam for some situation, you can use the notepad to write down a time interval for images to be deleted.
    b. The Notepad also has:
        i. Prepared statement to read to others who have questions or concerns.
        ii. List of places to deactivate the SenseCam.
        iii. Special instructions for private residences and workplace.
        iv. Experimenter contact info.
        v. Reminder to reactivate SenseCam.

Contact:

Please don't hesitate to contact me at any time if you have any questions or run into any problems.

~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
Jason R. Finley
Graduate Student, Department of Psychology
Cognitive Division
University of Illinois, Urbana-Champaign

Lab room: 275 Psychology Building

603 East Daniel Street
Champaign, IL 61820

Cell phone: [redacted for publication]
Lab phone: [redacted for publication]
Home phone: [redacted for publication]
Email: [redacted for publication]
Web: [redacted for publication]
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~

SenseCam Experiment 2, Summer 2010

Debriefing Sheet

Thank you for your participation in this study. There was no deception in this study. The procedures and purpose of the study are exactly as we described them in the initial information document and on your Consent Form.

The general purpose of this research is to develop a better understanding of the functioning of human autobiographical memory and to use this new knowledge to potentially improve the lives of individuals with disorders of memory. The specific aim of this experiment was to investigate the validity of memory for everyday experiences.

In most studies of peoples' memory for facts and experiences from their own personal lives (i.e., autobiographical memory), researchers have no way to verify the accuracy of information that participants recall. A few exceptions have been when researchers ask people to remember information for which there are objective records (e.g., grades they got in school, names of classmates, events of public record such as disasters). But the way that people remember those kinds of information may be different from the way people remember everyday experiences.

Some researchers have used daily diaries as records to cue and verify memory for everyday experience, but diary entries are usually made at the end of the day (by which time some forgetting may have already occurred), and the experiencer usually writes down only a small non-random selection of experience.

A further advancement in the study of autobiographical memory was the introduction of experience-sampling techniques, such as having participants write down aspects of their immediate experience whenever they are prompted (by a small device they carry) at random times throughout their day. However, such procedures necessarily interrupt experience, which may change the nature and/or memory of the experience.

Newer technology, such as the SenseCam, have now made it possible to capture objective records of certain aspects of experience in a much more continuous and inobstrusive way.

In this experiment, you wore a SenseCam for two days as you went about your normal everyday activities. Consistent with the initial information document and the Consent Form, no pictures were ever captured while the SenseCam was off or in Do Not Disturb mode, and all pictures falling in time periods marked for deletion were deleted without ever being seen. At delays of about 1 week and 1 month, you returned to the lab for a testing session in which we first had you remember as much as you could about various aspects of the original experience during a particular time period (Free Recall), then showed you SenseCam pictures from that time period, and finally gave you a chance to make corrections and additions to your Free Recall responses. The corrections that you made will give us an idea of how *accurate* your memory was for a given time period. The additions that you made will give us an idea of how *complete* your memory was for a given time period, and will also give us an idea of how much more information and what kinds of information are remembered when pictures of the original experience are presented as cues.

If you would like more information about SenseCam or about everyday/autobiographical memory, please consult the following references:

Brewer, W. F. (1988). Memory for randomly sampled autobiographical events. In U. Neisser & E. Winograd (Ed.), *Remembering reconsidered: Ecological and traditional approaches to the study of memory* (pp. 21-90). Cambridge: Cambridge University Press.

Cohen, G., & Conway, M. A. (2008). *Memory in the Real World* (3rd edition). Psychology Press.

Herry, E., Kapur, N., Williams, L., Hodges, S., Watson, P., Smyth, G., Srinivasan, J., Smith, R., Wilson, B. and Wood, K (2006). The use of a wearable camera, SenseCam, as a pictorial diary to improve autobiographical memory in a patient with limbic encephalitis, *Neuropsychological Rehabilitation*.

Steve Hodges, Lyndsay Williams, Emma Berry, Shahram Izadi, James Srinivasan, Alex Butler, Gavin Smyth, Narinder Kapur, and Ken Wood (2006) *SenseCam: a Retrospective Memory Aid*., In Dourish and A. Friday (Eds.): Ubicomp 2006, LNCS 4206, pp. 177 – 193, 2006. Springer-Verlag Berlin Heidelberg 2006. http://research.microsoft.com/~shodges/papers/SenseCam%20Ubicomp%202006%20(camera-ready).PDF

website: http://research.microsoft.com/en-us/um/cambridge/projects/sensecam/

website: http://www.viconrevue.com/

If you are interested in receiving an eventual writeup of the overall results of the study, just let us know (it will probably be done in about a year or two). At any point if you have further questions about this study, please contact Jason R. Finley, graduate student, Department of Psychology, University of Illinois, Champaign, IL 61820. Email: [redacted for publication]; cell phone: [redacted for publication]. Or contact the responsible project investigator: Professor William F. Brewer, Department of Psychology, University of Illinois, Champaign, IL 61820. Email: [redacted for publication]; phone: [redacted for publication]. If you have any questions concerning your rights as a participant in this experiment, you can contact the Subject Coordinator at [redacted for publication]. Once again, thank you very much for your participation in this experiment.