Running head: COMPARING PAPER AND SMARTPHONE DIARIES

Should participants be left to their own devices? Comparing paper and smartphone diaries in psychological research

Andrew Laughland and Lia Kvavilashvili

University of Hertfordshire, United Kingdom

Address for correspondence:

Lia Kvavilashvili

Department of Psychology and Sport Sciences

University of Hertfordshire

College Lane

Hatfield, Herts

AL10 9AB

United Kingdom

e-mail: l.kvavilashvili@herts.ac.uk

Word count for four introductions: 1517
Word count for the general discussion: 1512
Total word count including the abstract and the general audience summary: 7,692

Abstract

Growing smartphone ownership creates unprecedented opportunities for using participants' own smartphones as diaries to record transient phenomena in daily life. In three studies, we assessed the hypothesis that participant-owned smartphone diaries would result in superior compliance and higher number of recorded entries, than the traditional paper-diary method. Paper and smartphone diaries were compared for self-initiated recording of involuntary autobiographical memories (Studies 1 and 2), and everyday memory failures (Study 3). Diary-recording period (7-day, 1-day) was also examined by comparing results of Studies 1 and 2. Smartphone owners were highly compliant, carrying diaries, and making entries sooner. Nevertheless, significantly fewer memory events were recorded in smartphones than paper diaries in all studies. Moreover, the number of memories recorded in Study 2 (1-day) was significantly higher than recorded on day 1 of Study 1 (7-day), suggesting that shorter diary-keeping periods may be preferable. Implications and opportunities for improving smartphone-diary functionality are discussed.

Keywords*: Smartphone, paper diary, electronic diary, involuntary autobiographical memory, everyday memory failure*

General Audience Summary

In psychological research, capturing life as it is lived has typically been achieved by providing participants with diaries to record events of interest when they occur in everyday life. Initially, these were pen and paper based, but as technology has evolved, increasingly, electronic solutions have been employed. The last decade has seen an explosion in the ownership of smartphones. These are sophisticated devices which can be used to facilitate the capture of data on the go. Furthermore, owners are particularly diligent in carrying and consulting their smartphones at all times, while a common problem with paper diaries is that research participants forget to carry them, missing the opportunity to record events, or attempting to fill them in later with the potential for errors in recollection. For these reasons, making use of participant-owned smartphones in psychological research seems highly attractive, and many researchers are switching to this method. However, the assumptions that this will lead to increased compliance in carrying the diary and making entries, as well as improved data quality, have not been systematically tested. The purpose of the research described here was therefore to compare and contrast the old and new methods by randomly assigning smartphone owners to the paper- or smartphone-diary conditions. As demonstrated in the three studies reported in this paper, participants who used their own smartphones recorded substantially fewer events compared with those who were asked to carry paper diaries, although the nature of events recorded was comparable. We therefore advise researchers, intending to use apps on participants' smartphones, to exercise caution with the data they collect, particularly in respect of the rate, or frequency, of recording.

Should participants be left to their own devices? Comparing paper and smartphone diaries in psychological research

In psychological and medical research, participants are often asked to recall and rate events from their daily life retrospectively using questionnaire, survey, or interview methods. However, recalling events and rating their frequency or intensity after a delay may be negatively affected by guessing or personal beliefs (Cotter & Silvia, 2017; Schwarz, 2012). Recognising this recall bias, diary methods are considered the most appropriate tool for studying transient phenomena (e.g., intrusive thoughts, mood), which often go unnoticed, and are difficult to recall even after short delays (Bolger, Davis, & Rafaeli, 2003; Iida, Shrout, Laurenceau, & Bolger, 2012).

Some diary methods, known as Experience Sampling or Ecological Momentary Assessment (Shiffman, Stone, & Hufford, 2008), use random prompts to sample data. They are useful when participants can reasonably assess the phenomenon when prompted (e.g., mood, pain, current thoughts). Self-initiated diary recording is more appropriate when participants must complete diary entries at pre-set times (time-based recording) or in response to randomly occurring events/phenomena (event-based recording). This places additional demands on participants, since they must remember to carry a diary and record entries at particular times or when certain events occur (Takarangi, Garry, & Loftus, 2006).

Pen-and-paper diaries have been used for many years. However, from the 1990s, electronic diaries appeared, mainly as personal digital assistants (PDAs) programmed and lent to participants for the study. Electronic diaries were perceived superior to paper diaries, because they allowed prompting, time stamping, and enforcing and validation of entries, reducing retrospective back-filling of past events and forward filling of anticipated events. But there were also potential disadvantages such as keeping a device charged, technical failures, training burden, and cost.

To compare paper and PDA diaries directly, several studies within clinical and healthcare research examined participant compliance and the quality of the data collected with these two methods (for a review, see Dale & Hagen, 2007). In contrast, with the exception of a target paper by Green, Rafaeli, Bolger, Shrout, and Reis (2006), that elicited a considerable debate (Broderick & Stone, 2006; Takarangi et al., 2006; Tennen, Affleck, Coyne, Larsen, & DeLongis, 2006), there was little research on this topic in psychological literature. Moreover, a decade after this "paper or plastic" debate, smartphone technology (iPhones in 2007 and Android phones in 2008) resulted in PDAs being superseded by loaned, pre-programmed Android devices (Ainsworth et al., 2013; Palmier-Claus et al., 2012; Whalen, Odgers, Reed, & Henker, 2011).[1]

However, increased smartphone ownership worldwide creates an unprecedented opportunity to conduct psychological research using participants' *own smartphones* as electronic diaries. There are good reasons for using participants' own smartphones. First, participants purchase and learn to use the device themselves. Second, participants look after their smartphones and keep them charged, reducing technical problems. Third, smartphones have become constant companions (Clayton, Leshner, & Almond, 2015), with users unlocking their phones around 80 times per day (Statt, 2016). Consequently, we could expect very high compliance in carrying devices and completing entries soon after a to-be-recorded phenomenon is noticed. In contrast, a common problem with paper diaries is that participants forget to carry them, missing the opportunity to record events, or attempt to record later with the potential for errors in recollection.

Several researchers have already used participant-owned smartphone diaries to study mind-wandering, happiness and other phenomena (Killingsworth & Gilbert, 2010; MacKerron & Mourato, 2013; Monk, Heim, Qureshi, & Price, 2015). Nevertheless, before this new approach is fully adopted, it is necessary to (a) evaluate its effectiveness and comparability to traditional paper diaries on a number of critical dimensions (e.g.,

feasibility, compliance, data-completeness and equivalence), and (b) inform researchers about the best ways of collecting reliable and valid data while bearing in mind practical constraints (e.g., the length of recording period, burden to participants).

To address these important aims, we conducted three studies that systematically compared a paper diary and as near equivalent electronic diary app, installed on participants' own smartphones (iPhone or Android)[2], to study the frequency and nature of transient cognitive phenomena, using self-initiated event-contingent recording. In Studies 1 and 2, we compared paper and smartphone diaries using 7-day and 1-day recording periods, respectively, measuring the number and characteristics of recorded involuntary autobiographical memories (Berntsen, 2009). We predicted that the number of diary entries would be significantly higher in smartphone diaries, due to increased compliance carrying the device, compared with the paper diary. Contrary to predictions, significantly more diary entries were made in paper than smartphone diaries in both studies. To check the generalizability of these findings, in Study 3, participants recorded their everyday memory failures (Unsworth, Brewer, & Spillers, 2012). Additionally, we compared the diary recording periods of Study 1 (7-day) and Study 2 (1-day) in terms of the number and quality of recorded entries, to examine the optimum recording periods that minimise participant burden, but produce sufficient data.

Study 1

Involuntary autobiographical memories (IAMs) are memories of personal events that occur without any deliberate attempt to recall them (Berntsen, 2009; Schlagman & Kvavilashvili, 2008). They simply pop into mind, often in response to easily identifiable triggers (Mace, 2004; Mazzoni, Vannucci, & Batool, 2014) and during undemanding activities such as driving or washing up (Berntsen, 1996; Kvavilashvili & Mandler, 2004; Schlagman, Kliegel, Schulz, & Kvavilashvili, 2009).

Research on IAMs is one of the few areas in psychology that has been based predominantly on a diary method (Berntsen, 1996, 2009, 2010). Although several

laboratory methods have been developed (Berntsen, Staugaard, & Sørensen, 2013; Schlagman & Kvavilashvili, 2008), diaries are still used and will transition to smartphone-based studies. For example, Rasmussen, Ramsgaard, and Berntsen, (2015) have already used smartphone diaries to study IAMs, but they were loaned to participants. To our knowledge, no study has compared a participant-owned smartphone diary with a standard paper diary. This novel comparison is needed before research on IAMs and other spontaneous phenomena transitions to participant-owned smartphones.

In Study 1, participants were randomly assigned to a paper- or smartphone-diary conditions and had to complete a questionnaire every time they experienced an IAM. They were also allowed to acknowledge the memory if recording an IAM in full was not possible. The paper- and smartphone-diary conditions were compared in terms of the number of recorded entries, self-reported compliance, and several memory characteristics assessed in previous diary studies (Berntsen, 1996; Schlagman & Kvavilashvili, 2008).

We predicted that participants would carry their smartphones all the time in comparison to paper diaries. Therefore, they would record more IAMs, and possibly sooner after experiencing an IAM, than those in a paper-diary condition. However, if writing on paper was preferred over typing into smartphones, then memory descriptions would be shorter in the smartphone than paper diaries, or the number of acknowledged memories would be higher on the smartphone than in the paper diary. No differences between the two methods were expected for ratings of memory characteristics.

Method

Participants

The sample size calculations with the software G*Power (Faul, Erdfelder, Lang, & Buchner, 2007), were based on the assumption that there would be large group differences in the number of recorded IAMs, due to expected superior compliance rates in smartphone-diary participants. With the α-level set at .05, the statistical power (1 -

β) at .90, and aiming to detect large effects (η_p^2 = .16), the total required sample size was N = 58. Sixty participants, who owned an Apple iPhone or a smartphone with the Google Android operating system, were recruited from university students and staff. They were randomly allocated to smartphone-diary (*N* = 29, 26 female), and paper-diary (*N* = 31, 28 female) conditions.[3] In the smartphone-diary condition, there were 21 psychology and six non-psychology students and two staff. In the paper-diary condition, there were 24 psychology and five non-psychology students and two staff. The mean age of the smartphone group was 24.14 years (*SD* = 8.16, range 18-51), and did not differ from the mean age of 24.71 years (*SD* = 9.28, range 18-51) in the paper-diary group (*F* < 1). The iPhone was more common than Android in both groups, but the proportion of iPhone and Android ownership did not differ by condition, χ^2 (1, *N* = 60) = 2.55, *p* = .11. There were no group differences in the length of ownership of a smartphone with a mean of 2.01 years (*SD* = 1.09) in the smartphone group, and 2.07 years (*SD* = 1.72), in the paper diary group (*F* < 1).

Study Design

We used a one factor ANOVA with the diary condition (paper vs. smartphone) as a between subjects variable. The main dependent variables involved the number of recorded diary entries (full and acknowledged), and self-reported compliance, which was measured by (a) the number of days participants reported keeping the diary with them, (b) the percentage of IAMs that they were able to record and acknowledge (out of all the IAMs experienced during the recording period), (c) self-reported delay between the IAM occurrence and its recording, and (d) the number of words in memory descriptions. The ratings of IAMs on several dimensions (e.g., vividness, pleasantness, previous rehearsal, specificity of memories) were also examined.

Materials

Smartphone Usage Questionnaire. This questionnaire was completed before keeping the diary. It contained several questions about participants' use of smartphones (e.g., for voice calls, texting, emails), and technical skills, to ensure the comparability between the two conditions on these measures (see Supplementary Materials).

Diary Compliance Questionnaire. This questionnaire was completed after the diary-keeping phase. Participants indicated whether they carried the diary with them every day of the study (*yes/no*), and if the "no" option was chosen, on how many days they forgot. They were also asked to estimate the percentage of IAMs they were able to record out of all IAMs experienced in the 7-day period. Finally, they rated how easy they found (a) keeping the diary with them at all times, and (b) recording their memories in the diary (*Very easy, Somewhat easy, Somewhat difficult, Very difficult*).

Paper Diary and the Smartphone Diary App. In the paper-diary condition, participants received an A5 paper diary booklet containing 32 identical pages, with 11 questions per page to be completed for each IAM experienced (for an example diary page see Appendix 1). The smartphone app was built by the first author. It was designed to be as similar as possible to the paper diary both in terms of simplicity of completing a diary entry and the questionnaire format/interface. As with the paper diary, there was no prompting, and any field could be left unanswered without participants being alerted to the omission at the point of submitting a diary entry. The app was available as an icon on the home screen of the phone. Data were kept locally on the phone for the seven days so there was no need for an Internet connection. This was important as it meant entries could be made at any time, even out of signal.

Procedure

The study was advertised as a 7-day diary study of IAMs. Participants had to own a smartphone and be willing to keep a diary of their IAMs using either a paper diary or their own smartphone diary app. They were not told that the purpose was to

compare the two modes of data collection. Following consent, participants supplied demographic information, and completed the Smartphone Usage Questionnaire. Participants in the paper-diary condition were given the diary booklet. In the smartphone-diary condition, the researcher installed the app on participants' smartphones while they completed questionnaires.

The researcher briefed each participant individually for 20-30 min, taking them through each item on the diary page, or on the smartphone screen. The concept of IAMs was carefully explained, including how they differ from voluntary autobiographical memories, by providing relevant examples. A clear distinction between specific memories of single, one-time events (e.g., *an argument with a teacher*) and general memories of repeated events (e.g., *walking to school every morning*) or extended events (e.g., *a weekend in Paris*) was also illustrated with relevant examples.

Participants were asked to record any IAMs that occurred over the next seven days, starting from waking the day after the briefing, so that only full days were recorded. In addition to verbal briefing, they were given written instructions on how to complete their paper or smartphone diary. Participants were urged to keep the diary with them at all times and record their IAMs immediately, or as soon as possible after occurrence. If they could not complete the diary entry immediately (e.g., they were attending a meeting, or driving) and later felt that they could not remember the key characteristics (e.g., triggers, vividness), they could record the IAM as a tick (in a table on the inside front cover of the paper diary) or by pressing an acknowledge memory button in the app. "Acknowledge" button presses in the app were time-stamped. No expectations were set regarding the number of memories that could be recorded. Participants were informed that they could have very few or many. If for some time they did not experience any memories at all, that was fine, too. The most important thing was that they recorded only genuinely involuntary memories and did not try to recall them deliberately even when they had not recorded any memories for some time.

Finally, each participant made an appointment to return one, or two days after the 7-day diary-keeping period, to hand back the paper diary or have the electronic diary data uploaded to a data-server. At this time, they completed the Diary Compliance Questionnaire and were debriefed.

Results and Discussion

Both parametric and non-parametric methods of analysis were used depending on the type of the dependent variable used. The α-level was set at .05, and the effect size, measured by partial eta-squared (η_p^2), was defined as .01, .06, and .16 for small, medium and large effects, respectively (Cohen, 1988). Before analysing the data, the equivalence of paper- and smartphone-diary conditions was established by showing no statistically significant differences in participants' use of their smartphones, self-rated typing ability, or adaptability to technology (see Table 1 in Supplementary Materials).

Measures of Compliance in Paper- and Smartphone-diary Conditions

To address the hypothesis that compliance rates would be significantly higher in the smartphone- than paper-diary condition, analyses were conducted on several different measures of compliance and the length of memory descriptions. The mean number of words in memory descriptions in the smartphone-diary condition (M = 14.81, SD = 10.01) was not significantly different from the mean number of words in the paper-diary condition (M = 13.69, SD = 5.60), $F < 1$.

Participants' responses in the Diary Compliance Questionnaire showed that there was no significant difference between the groups in the percentage of IAMs that they reported they were able to record and acknowledge (69% in the smartphone-diary, 68% in paper-diary condition, $F < 1$), out of all the IAMs they had over the 7-day period. However, groups differed significantly on the remaining questions. In the smartphone group, 79% reported that carrying the diary with them at all times was "very easy", while only 21% of participants in the paper-diary condition chose this option on a 4-point rating scale (*very easy, somewhat easy, somewhat difficult, very difficult*), χ^2

(3, $N = 60$) = 21.94, $p < .0001$. Similarly, 68% of participants in the smartphone-diary condition reported that recording memories in the diary was "very easy" in comparison to only 32% of participants in the paper-diary condition, χ^2 (3, $N = 60$) = 8.54, $p = .036$.

Moreover, while none of the smartphone-diary participants reported forgetting to carry their smartphone with them on any of the seven days of the study, 35% of paper-diary participants reported forgetting to carry the diary for one (19%), two (10%) or even three days (6%). In a one-sample *t*-test, the mean number of forgotten days in the paper-diary condition was significantly different from zero, $t(30) = 3.50$, $p = .001$.

To examine the speed of recording memories in the diary, we calculated the proportion of memories recorded by each participant within 10 min of their reported occurrence. The mean proportion of recorded IAMs in the 10-min window was significantly higher in the smartphone-diary ($M = .69$, $SD = .25$) than in the paper-diary condition ($M = .52$, $SD = .29$), $F(1,58) = 5.65$, $p = .021$, $\eta_p^2 = .067$, indicating more prompt recording on the smartphone. Taken together, these results provide strong support for our prediction that participants in the smartphone-diary condition would exhibit significantly better compliance than those in the paper-diary condition.

The Number of Recorded Involuntary Memories

All participants kept a diary and fully recorded at least two IAMs over the 7-day period.[4] In total, 835 memories were fully recorded by completing a diary questionnaire (559 in the paper- and 276 in the smartphone-diary condition), and 442 memories were acknowledged by ticking a grid on the inner cover page of the paper-diary (304 in the paper-diary) or pressing a button in the app (138 in the smartphone-diary condition).[5]

In line with previous research, there was a large variability in the number of recorded IAMs in both conditions (see Table 1). To normalise the positively skewed data, the number of fully recorded and acknowledged IAMs were square root transformed before submitting them to one-way between subjects ANOVA. Contrary to our predictions, results showed that participants recorded almost twice as many entries

in the paper- than smartphone-diary condition, $F(1,58) = 16.74$, $p = .0001$, $\eta_p^2 = .22$. Similar results were obtained for the number of acknowledged memories, albeit with smaller effect size, $F(1,58) = 4.15$, $p = .046$, $\eta_p^2 = .067$.

We also examined the number of fully recorded IAMs across the seven days of diary-keeping to see if different patterns emerged in the two conditions. The results of a 2 (condition) x 7 (days) mixed ANOVA on the number of recorded IAMs (square root transformed) showed that the interaction between the condition and days was not statistically significant, $F(6,348) = 1.91$, $p = .078$, $\eta_p^2 = .03$. However, there was a significant main effect of days, $F(6,348) = 6.48$, $p < .0001$, $\eta_p^2 = .10$, in addition to the main effect of condition (see Figure 1). Follow-up post-hoc tests, with Bonferroni correction for multiple comparisons, indicated that the mean number of IAMs reported on Day 1 ($M = 1.53$, $SD = 0.69$) was not significantly different from IAMs reported on Day 2 ($M = 1.28$, SD = 0.78), and Day 3 ($M = 1.20$, $SD = 0.72$) ($p = .13$, and $p = 0.06$, respectively). However, it was significantly higher than the mean number of IAMs recorded on Day 4 ($M = 1.05$, $SD = 0.83$), Day 5 ($M = 0.97$, $SD = 0.70$), Day 6 ($M = 1.08$, $SD = 0.80$) and Day 7 ($M = 1.14$, $SD = 0.78$) (all $p_s < .009$). None of the other comparisons were significant ($p_s > .17$).

A similar 2 (condition) x 7 (days) mixed ANOVA on the number of acknowledged memories resulted only in a significant main effect of condition ($F(1,58) = 4.47$, $p = .039$, $\eta_p^2 = .07$), but no significant effect of days ($F < 1$) or the condition by days interaction ($F(6,348) = 1.99$, $p = .067$, $\eta_p^2 = .03$).

Comparing Paper- and Smartphone-diary Conditions on Other Variables

For each participant, we calculated the mean proportion of specific IAMs and the mean ratings of vividness (on a 7-point rating scale), pleasantness and rehearsal (on 5-point rating scales). These means were entered into one-way between-subjects ANOVAs (see Table 2). No significant differences between the paper- and smartphone-

diary conditions were obtained. The two conditions also did not differ in the mean concentration ratings made on a 5-point scale (smartphone diary $M = 3.21$, $SD = 0.75$; paper diary $M = 3.05$, $SD = 0.58$) ($F < 1$). Finally, for each participant, we also calculated the proportion of memories reported to have an internal trigger, external trigger or no trigger. The means of these proportions as a function of condition and type of trigger are presented in Table 3. A 2 (condition) by 3 (trigger type) mixed ANOVA resulted in the significant main effect of trigger type ($F(2,116) = 145.49$, $p < .00001$, $\eta_p^2 = .71$) with the mean proportion of external triggers ($M = 0.71$, $SD = 0.20$) being significantly higher than the mean proportions for internal triggers ($M = 0.17$, $SD = 0.15$) and no triggers ($M = 0.12$, $SD = 0.16$) ($p_s < .001$), which did not differ from each other ($p = .15$). The condition by trigger type interaction was not significant ($F < 1.01$).

In summary, although participants in the smartphone-diary condition displayed superior compliance on several measures, they recorded significantly fewer IAMs than those in the paper-diary condition. However, the groups did not differ in the nature of reported memory characteristics or conditions in which they were reported to occur. In addition, the data presented in Tables 2 and 3 are fairly similar to what has been reported in previous paper diary studies of IAMs (Berntsen, 1996, 1998; Berntsen & Hall, 2004; Schlagman & Kvavilashvili, 2008).

Study 2

In Study 2, we wanted to examine if the superiority of paper diaries, in terms of the number of recorded IAMs, would be replicated using a much shorter, 1-day, recording period. Reducing recording to one day was interesting for two reasons. First, it minimised the chances that paper-diary participants would forget to keep a diary, making compliance levels more comparable across the two conditions. Second, it enabled us to examine whether the extended periods of recording, often used in diary studies, are needed. If participants are more aware of being in a study and more willing to monitor their memories for a shorter period, as demonstrated by recent 1-day studies,

in which participants acknowledged a large number of IAMs with mechanical clickers (Rasmussen & Berntsen, 2011; Rasmussen et al., 2015), then the number of IAMs recorded in a 1-day diary of Study 2 would be greater than on Day 1 of the 7-day diary in Study 1.

Method

Participants

A priori power analysis using G*Power, based on the large effect obtained for the number or recorded IAMs in Study 1 (η_p^2 =.22), with the α-level set at .05, and the statistical power (1 - β) at .90, resulted in the total required sample size of N = 40. We recruited 49 participants (all but one were psychology students), who owned a smartphone, and were randomly allocated to two conditions: smartphone diary (*N*=23, 19 female), and paper diary (*N*=26, 21 female). The mean age of the smartphone group was 21.74 years (*SD* = 4.64, range 18-36), and did not differ significantly from the mean age of 23.19 years (*SD* = 7.61, range 18-51) in the paper-diary group ($F < 1$). No significant differences between the conditions were obtained in terms of participants' smartphone usage and self-rated technical ability (all $p_s > .10$).

Materials and Procedure

The paper diary (with 32 pages) and the smartphone app were identical to those used in Study 1. The instructions and the procedure were also the same, except that participants were briefed on Day 1, kept the diary only on Day 2 and returned on Day 3.

Results and Discussion

Measures of Compliance in Paper- and Smartphone-diary conditions

The mean number of words used for memory descriptions in the smartphone-diary condition (*M* = 10.70, *SD* = 6.67) was nominally lower than the mean number of words in the paper-diary condition (*M* = 14.77, *SD* = 7.59). However, the difference was not statistically significant, $F(1,47) = 3.93$, $p = .053$, $\eta_p^2 = .08$, which could potentially indicate the lack of power, given the medium size effect.

The groups also did not differ significantly in terms of the percentage of IAMs they reported being able to record and acknowledge (62% in the smartphone-diary, 66% in paper-diary condition), out of all the IAMs experienced during the day. However, as in Study 1, 70% of the smartphone-diary participants said that keeping the diary with them was "very easy", in comparison to just 16% in the paper-diary condition, χ^2 (3, N = 49) = 14.48, p = .002. Similarly, 70% of participants in the smartphone-diary condition said that recording memories was "very easy", compared with 42% in the paper-diary condition, but this difference was not significant, χ^2 (3, N = 49) = 5.31, p = .15. As the diary was kept for one day only, there were no instances of forgetting to carry it.

Unlike Study 1, the proportion of IAMs recorded within 10 min of their reported occurrence, was not significantly higher in the smartphone- (M = 0.73, SD = 0.29), than paper-diary condition (M = 0.76, SD = 0.27) ($F < 1$). Overall, these results show that the superiority of smartphones over paper diaries in compliance rates obtained in Study 1 is much diminished with 1-day recording, providing support for the idea that participants are more on-task during shorter recording periods.

The number of recorded memories

All participants fully recorded at least one IAM. A total of 285 memories were fully recorded (196 in the paper diary, and 89 in the smartphone diary), and 75 were acknowledged (30 in the paper diary and 45 in the smartphone). The mean numbers of recorded and acknowledged IAMs in both conditions are presented in the middle panel of Table 1. The analysis of variance, carried out on square-root transformed data, showed that participants in the paper-diary condition fully recorded almost twice as many memories than those in the smartphone-diary condition, F(1,47) = 10.21, p = .002, η_p^2 = .18. However, there was no significant difference for the number of acknowledged memories, F(1,47) = 2.82, p = .10, η_p^2 = .057.

Comparing Paper- and Smartphone-diary Conditions on Other Variables

The mean ratings of memory characteristics and results of one-way ANOVAs presented in Table 2 show that there were no significant differences between the two conditions in terms of ratings of vividness and rehearsal or the proportion of specific IAMs recorded. However, the ratings of memory pleasantness were higher in the paper- than smartphone-diary condition. The mean ratings of concentration were nominally higher in the smartphone- (M = 3.13, SD = 0.66) than paper-diary condition (M = 2.74, SD = 0.73), but the difference was not statistically significant, $F(1,47) = 3.90$, $p = .054$, $\eta_p^2 = .08$. Finally, as in Study 1, there was a significant main effect of trigger type ($F(1.58,74.21) = 156.67$, $p < .00001$, $\eta_p^2 = .77$), with the majority of IAMs being reported to have occurred in response to external triggers (see Table 3). The condition by trigger type interaction was not significant (F < 1.44).

Number of Recorded Memories in Day 1 (Study 1 versus Study 2)

Finally, we compared the fully recorded and acknowledged memories in this 1-day study with the number of memories recorded in Day 1 of the 7-day study (Study 1). A 2 diary period (1-day vs. 7-day) x 2 diary type (paper vs. smartphone) between subjects ANOVA on the number of fully recorded memories (square root transformed) resulted in a significant main effect of diary period, $F(1,105) = 26.88$, $p < .0001$, $\eta_p^2 = .20$. Overall, significantly more IAMs were fully recorded in 1-day diary in Study 2 (M = 5.82, SD = 4.71) than on Day 1 of a 7-day diary in Study 1 (M = 2.82, SD = 2.55). This main effect did not interact with diary type, $F(1,105) = 2.42$, $p = .12$, $\eta_p^2 = .02$. As expected, the main effect of diary type was also significant ($F(1,105) = 11.87$, $p = .001$, $\eta_p^2 = .10$) with more IAMs recorded in the paper- (M = 5.23, SD = 4.71) than smartphone-diary condition (M = 3.00, SD = 2.28). A similar 2 (diary period) by 2 (diary type) ANOVA on the mean number of acknowledged memories did not result in any significant main or interaction effects (all Fs < 2.38).

Overall, Study 2 resulted in several important findings. First, reducing diary keeping to one day, improved compliance rates in the paper diary condition, Indeed, no participant forgot to keep the diary, and the mean proportions of IAMs, recorded within 10 min of having a memory, were not statistically different in paper- and smartphone-diary conditions. Second, the number of fully recorded memories in the paper-diary condition was still significantly higher than in the smartphone-diary condition, replicating the main finding of Study 1. Third, although the results of cross-study comparisons should be interpreted with caution, the number of fully recorded IAMs in the 1-day diary of Study 2 was significantly higher than the number recorded on Day 1 of the 7-day diary in Study 1. This is the first direct evidence showing that reducing the diary-keeping period does not proportionately reduce the number of recorded entries.

Study 3

In Study 3, we tested the generalizability of results of Studies 1 and 2 by asking participants to keep a diary of their everyday memory failures (EMFs), which include a variety of retrospective, prospective and absent-minded errors (e.g., forgetting someone's name, missing an appointment, brushing teeth with a shaving cream) (Kvavilashvili, Kornbrot, Mash, Cockburn, & Milne, 2009; Unsworth et al., 2012). EMFs may cause annoyance and have negative consequences (McDaniel & Einstein, 2007; Reason & Mycielska, 1982). Having an embarrassing EMF may even remind the participant that they are in a diary study. Therefore, EMFs should be noticed more easily than IAMs (e.g., Mace, Bernas, & Clevinger, 2015). It is possible that these features of EMFs will improve recording rates in smartphone diaries and reduce or eliminate differences in the two diary conditions obtained in Studies 1 and 2.

Since EMFs occur less frequently than IAMs, the few diary studies, reported in the literature, have used recording periods of several weeks (Heine, Ober, & Shenaut, 1999; Reason & Mycielska, 1982; Terry, 1988; but see Unsworth et al., 2012). In Study 3, participants kept a diary of EMFs for seven days to be comparable with 7-day

IAM diaries used in Study 1, and to determine whether shorter 7-day recording periods still provided enough entries.

Method

Participants

Based on large effects obtained for the number or recorded IAMs in Study 1 (η_p^2= .22) and Study 2 (η_p^2 = .18), a priori power analysis using G*Power, resulted in the total required sample size of N = 38 (with the α-level set at .05, the power (1 - β) at .85, and the expected effect size η_p^2 =.20). Thirty-seven participants (university students, staff and alumni) were randomly allocated to smartphone-diary (*N* = 19, 16 female), and paper-diary (*N* = 18, 9 female) conditions. In the former, there were 16 psychology students and three staff. In the latter, there were 12 psychology students, 1 staff and 5 alumni. Participants' mean age did not differ in the smartphone- (M = 21.79, *SD* = 7.03, range 18-46) and paper-diary (*M* = 26.28, *SD* = 9.87, range 18-49) conditions, $F(1,35) = 2.56$, $p = .12$, $\eta_p^2 = .07$.

Materials and Procedure

The paper and smartphone diaries were identical in format to diaries used in Studies 1 and 2, except that questions about IAMs were replaced by questions about EMFs (see Appendix 2). The procedure and instructions were identical to those used in Study 1 and 2, but we briefed participants for 20-30 min about EMFs with examples of retrospective, prospective, and absent-minded errors (see Kvavilashvili et al., 2009). Participants had to record any memory failure that occurred over the seven days, starting from the day after the briefing.

Results and Discussion

Measures of Compliance in Paper- and Smartphone-diary conditions

The analyses of the Diary Compliance Questionnaire showed that the two groups did not differ reliably in the percentage of EMFs they reported being able to record and acknowledge in the diary (83% in the smartphone-, 82% in paper-diary conditions, *F* <

1). However, 95% of the smartphone-diary participants said keeping the diary with them was “very easy”, compared to just 33% in the paper-diary condition, χ^2 (4, N = 37) = 15.99, p = .003. Similarly, 79% of smartphone-diary participants said recording EMFs in the diary was “very easy”, compared to 62% in the paper-diary condition, but this difference was not significant, χ^2 (3, N = 37) = 1.46, p = .69. While no smartphone-diary participant forgot to carry their device, 33% of the paper-diary participants forgot to carry their diary on at least one day of the seven, and the mean number of forgotten days was significantly different from zero, t(17) = 2.72, p = .008. Finally, although the mean proportion of EMFs recorded within 10 min of their occurrence was nominally higher in the smartphone- (M = 0.58, SD = 0.33) than paper-diary condition (M = 0.45, SD = 0.32), this difference was not statistically significant F(1,35) = 1.42, p = .24, η_p^2 =.04.

The Number and Characteristics of Recorded Memory Failures

All participants recorded at least two EMFs. The mean numbers of fully recorded and acknowledged EMFs are presented in a lower panel of Table 3. The analysis of variance, carried out on square-root transformed means, showed that participants in the paper-diary condition fully recorded almost twice as many EMFs than those in the smartphone-diary condition, F(1,35) = 4.57, p = .04, η_p^2 = .12. Unlike Studies 1 and 2, there were very few acknowledged memory failures. There was one outlier with 33 acknowledged EMFs in the paper-diary condition, and one with eight acknowledged EMFs in the smartphone-diary condition. When these were excluded, the mean numbers of acknowledged EMFs in the two conditions were not significantly different, F(1,33) = 1.31, p = .26., η_p^2 = .04. There were also no significant differences between the conditions for mean ratings on any of the four scales used (mood before, relaxed or stress level before, seriousness of the lapse, and how upset one was by the EMF), all Fs < 1.

In summary, Study 3 replicated the results of Studies 1 and 2, showing that the superiority of paper diaries, in terms of the number of fully recorded entries, was

maintained, even though participants had to record EMFs, which are noticeably different from IAMs. The differences between the phenomena were obvious from the findings showing that fewer EMFs were recorded over the 7-day period than IAMs in Study 1 (see Table 1), and that there were very few acknowledged entries. Although EMFs occurred less frequently, they were probably more noticeable and more memorable than IAMs, and participants were able to fully record them even after delays (as reflected by reduced proportions of entries made within 10 min when compared to the data from Study 1).

General Discussion

Three main findings emerged. First, smartphone-diary participants displayed significantly better self-rated compliance across seven days, keeping the diary with them at all times (Studies 1 and 3) and recording entries significantly sooner than those in the paper-diary condition (Study 1). Second, significantly fewer IAMs were recorded fully by participants on their smartphones, compared with paper diaries (Studies 1-2). This effect was replicated in Study 3 on EMFs, showing that the finding is not unique to IAMs. The results for the number of acknowledged entries were less consistent, with no significant differences between the diary conditions in Studies 2 and 3. Third, reducing diary keeping to one day (Study 2), did not proportionately decrease the number of entries recorded in paper or smartphone diaries. These findings have important implications for research using diaries for data collection.

Compliance rates in paper- versus smartphone-diaries

Superior compliance with smartphones is not surprising, given that people carry their smartphones everywhere and frequently consult them. Participants estimated the proportion of recorded events, the number of days they forgot to carry their diaries, and rated the ease of carrying the diary at all times. Findings concerning these estimates were consistent across studies, which is reassuring, since paper-diary users appeared

honest in admitting to forget keeping the diary some days, while smartphone users were confident they had not.

It is difficult to assess actual compliance with paper diaries, without using elaborate technology (Stone, Shiffman, Schwartz, Broderick, & Hufford, 2003), but these self-report measures probably gave a relatively accurate view of compliance, since participants were free to record, or not, in their diaries without consequence. This freedom was further enhanced by allowing participants the simpler acknowledge option. Moreover, high compliance rates reported in smartphone-diary conditions are in line with the majority of studies reviewed by Dale and Hagen (2007), who found increased compliance with loaned PDA diaries over paper, and sometimes, greater participant preference.

Number of Diary Entries in Paper versus Smartphone Diaries

The key finding was that in all three studies, fewer entries were recorded in participant-owned smartphones than paper diaries. This effect cannot be explained by paper diaries being easier to use, because every effort was made to match the smartphone app to the paper diary in the appearance and functionality. Completing the smartphone-diary entry took between 1-2 min (similar to what participants have reported informally in our previous paper-diary studies), and memory descriptions were of comparable length in both conditions.

The only potential difference between the diaries was that paper diaries had 32 pages, while the number of entries in the smartphone app was unlimited. However, all participants were explicitly instructed that we had no expectations about how many entries they would record. Moreover, smartphone-diary participants could record any number of entries, while paper-diary participants could have felt more constrained by the finite number of diary pages, yet they recorded more entries.

The difference between the conditions is probably due to some other processes. One possibility is that paper diaries acted as effective incidental cues to remind

participants that they were in the diary study. Indeed, participants could easily forget that they were supposed to be monitoring for and recording a certain phenomenon. However, seeing the paper diary throughout the day may have increased participants' awareness of being in the study and renewed, or reinforced, their monitoring, as reported informally by some paper-diary participants (*cf.* Kvavilashvili & Fisher, 2007). Moreover, if participants were unable to make a recording, seeing a paper diary later could have also reminded them to complete an entry.

In contrast, the smartphone users probably recorded fewer instances because the app did not stand out from other apps, and their smartphones were highly familiar devices serving multiple functions, in addition to recording memories. Consequently, smartphones were less likely to remind participants that they were in a memory study, in the way that the alien paper diary did, with its one, out-of-the-ordinary, purpose. The inability of the smartphone diary to act as an effective reminder is similar to the cue overload and fan effects in episodic memory (Anderson & Reder, 1999; Watkins & Watkins, 1975; see also Berntsen et al., 2013).

Although paper diaries could both increase general monitoring of the phenomenon under investigation, and remind participants to complete a missed entry after a delay, results of Study 2 showed that 1-day paper-diary participants recorded their entries within 10 min as often as smartphone-diary participants did. This would reduce chances for the paper diary to act as a reminder for delayed or missed recordings. Therefore, the more likely explanation for increased entries in paper-diaries is that seeing paper diaries enhanced participants' awareness of being in the study. It is also possible that smartphone participants got distracted with notifications from other apps when they picked up their phone to record a memory. This distraction would not occur if participants had been lent a smartphone to use as an electronic diary. Future research should clearly evaluate these alternative (albeit not mutually exclusive) explanations.

The Length of Recording Period (1 Day versus 7 Days)

The idea that participants may struggle to maintain the awareness of being in the study, especially over extended periods, is supported by comparing the number of recorded entries in 7- and 1-day diaries of Studies 1 and 2. Indeed, the analysis of diary entries across the seven days of Study 1 suggests that there was an initial enthusiasm, or effort, to record IAMs, but this fell, and significantly so from Day 4 onwards. Moreover, the number of recorded IAMs in 1-day paper and smartphone diaries of Study 2 constituted 42% and 41%, respectively, of the total number of IAMs recorded in the 7-day diaries of Study 1. The finding of higher rates of recorded entries in reduced recording periods of Study 2 is consistent with some IAM studies where participants recorded memories for periods shorter than in standard one or two-week diary studies (Kamiya, 2013; Rasmussen et al., 2015). Together, results strongly suggest that diary-keeping periods of several days, or weeks, currently used by researchers of self-observed cognitive phenomena could be reduced (e.g., to just a few days, or even one day, depending on the aims of the study and the phenomenon under investigation). [6]

Contribution to Research Practice

The findings reported here are important and timely, because the use of smartphones in psychological, social science, and clinical research will continue to increase, as anticipated by the Smartphone Manifesto by Miller (2012). However, researchers must balance the convenience to participants, and themselves, of electronic data gathering on devices supplied by participants, with the limitations of the approach demonstrated in the present studies. While participants carried their smartphones with them consistently, the number of entries recorded in the app were disappointing. Consequently, we need to adapt research methods, briefings and prompts to mitigate issues of logging fewer entries of studied phenomena in the smartphone diaries.

We do not want to dismiss the opportunity of using participant-owned devices for recording event-contingent phenomena, which are not amenable to experience

sampling methods. But we caution that the anticipated benefits might need to be supported with some means of raising participants' awareness to monitor. We recommend further work to ensure that the smartphone diary reminds participants that they are in a diary study. This could take the form of technological solution such as a daily, or more frequent, text message, or other form of smartphone alert. Alternatively, participants may be asked to change the usual appearance of their phone (e.g., with a coloured phone case, or modified screen image) to mimic the incidental cueing aspect of the paper diary.

The paper diary has received an unexpected boost to its reputation in our studies. While increasingly seen as unacceptable in medical and regulated diary studies (Coons et al., 2014), paper diaries still seem appropriate for studying transient phenomena where self-monitoring and self-initiated recording is required, with no pressure to deliver a certain number of entries, which reduces the chances of fabricating recorded events. However, to further test the generalizability of our findings, future studies should compare paper and smartphone diaries using different spontaneous phenomena (e.g., musical earworms, intrusive memories) and participant groups (e.g., older adults).

Finally, reducing the diary-keeping period appears to be justified in terms of the quantity and quality of data collected, and because it reduces the burden on both participants and researchers. It allows more participants to be processed and increases the likelihood of recruitment. A 3-day diary, with a briefing on Monday and debriefing on Friday, appears optimal (although participants' work patterns may bias the phenomena, or the compliance), but a 1-day paper, or smartphone diary (with appropriate attention to cueing) may also be acceptable, especially when studying phenomena that occur frequently in everyday life.

In conclusion, with online recruitment, and increasingly sophisticated data gathering and psychometric tools moving to the smartphone (Thai & Page-Gould, in press), we are in no doubt that participant-owned smartphone diaries will become the

standard tool, and one which participants will largely prefer. However, these studies serve as a timely reminder that each new generation of technology brings its challenges.

Author contributions

AL and LK jointly generated the idea for the studies and developed study materials. AL developed the smartphone app and collected the data. AL analyzed the data, and LK verified the accuracy of those analyses. AL wrote the first draft of the manuscript, and both authors critically edited it. Both authors approved the final submitted version of the manuscript.

Acknowledgments

We would like to acknowledge the assistance of Vida Zohrabi Kia and Dimitrios Nahkid for collecting some of the data in Study 1 and Study 2, respectively. In addition, we would like to thank John Done for helpful conversations, which contributed to the design of this research.

References

Ainsworth, J., Palmier-Claus, J. E., Machin, M., Barrowclough, C., Dunn, G., Rogers, A., … Lewis, S. (2013). A comparison of two delivery modalities of a mobile phone-based assessment for serious mental illness: Native smartphone application vs text-messaging only implementations. *Journal of Medical Internet Research*, *15*, e60. https://doi.org/10.2196/jmir.2328

Anderson, J. R., & Reder, L. M. (1999). The fan effect: New results and new theories. *Journal of Experimental Psychology: General*, *128*, 186–197. https://doi.org/10.1037/0096-3445.128.2.186

Berntsen, D. (1996). Involuntary autobiographical memories. *Applied Cognitive Psychology*, *10*, 435–454. https://doi.org/10.1002/(SICI)1099-0720(199610)10:5<435::AID-ACP408>3.0.CO;2-L

Berntsen, D. (1998). Voluntary and involuntary access to autobiographical memory. *Memory*, *6*, 113–141.

Berntsen, D. (2009). *Involuntary autobiographical memories: An introduction to the unbidden past*. Cambridge, UK ; New York: Cambridge University Press.

Berntsen, D. (2010). The unbidden past: Involuntary autobiographical memories as a basic mode of remembering. *Current Directions in Psychological Science*, *19*, 138–142. https://doi.org/10.1177/0963721410370301

Berntsen, D., & Hall, N. M. (2004). The episodic nature of involuntary autobiographical memories. *Memory & Cognition*, *32*, 789–803.

Berntsen, D., Staugaard, S. R., & Sørensen, L. M. T. (2013). Why am I remembering this now? Predicting the occurrence of involuntary (spontaneous) episodic memories. *Journal of Experimental Psychology: General*, *142*, 426-444.

Bolger, N., Davis, A., & Rafaeli, E. (2003). Diary methods: Capturing life as it is lived. *Annual Review of Psychology*, *54*, 579–616. https://doi.org/10.1146/annurev.psych.54.101601.145030

Broderick, J. E., & Stone, A. A. (2006). Paper and electronic diaries: Too early for conclusions on compliance rates and their effects: Comment on Green, Rafaeli, Bolger, Shrout, and Reis (2006). *Psychological Methods*, *11*, 106–111. https://doi.org/10.1037/1082-989X.11.1.106

Clayton, R. B., Leshner, G., & Almond, A. (2015). The extended iSelf: The impact of iPhone separation on cognition, emotion, and physiology. *Journal of Computer-Mediated Communication*, *20*, 119-135. https://doi.org/10.1111/jcc4.12109

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2 edition). Hillsdale, N.J: Routledge.

Coons, S. J., Eremenco, S., Lundy, J. J., O'Donohoe, P., O'Gorman, H., & Malizia, W. (2014). Capturing patient-reported outcome (PRO) data electronically: The past, present, and promise of ePRO measurement in clinical trials. *Patient, 8, 301-309*. https://doi.org/10.1007/s40271-014-0090-z

Cotter, K. N., & Silvia, P. J. (2017). Measuring mental music: Comparing retrospective and experience sampling methods for assessing musical imagery. *Psychology of Aesthetics, Creativity, and the Arts*, *11*, 335–343. https://doi.org/10.1037/aca0000124

Dale, O., & Hagen, K. B. (2007). Despite technical problems personal digital assistants outperform pen and paper when collecting patient diary data. *Journal of Clinical Epidemiology*, *60*, 8–17. https://doi.org/10.1016/j.jclinepi.2006.04.005

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175–191. https://doi.org/10.3758/BF03193146

Green, A. S., Rafaeli, E., Bolger, N., Shrout, P. E., & Reis, H. T. (2006). Paper or plastic? Data equivalence in paper and electronic diaries. *Psychological Methods*, *11*, 87–105. https://doi.org/10.1037/1082-989X.11.1.87

Heine, M. K., Ober, B. A., & Shenaut, G. K. (1999). Naturally occurring and experimentally induced tip-of-the-tongue experiences in three adult age groups. *Psychology and Aging*, *14*, 445.

Iida, M., Shrout, P., Laurenceau, J., & Bolger, N. (2012). Using diary methods in psychological research. In *APA handbook of research methods in psychology: Foundations, planning, measures and psychometrics*. American Psychological Association.

Kamiya, S. (2013). Relationship between frequency of involuntary autobiographical memories and cognitive failure. *Memory*, 1–13. https://doi.org/10.1080/09658211.2013.838630

Killingsworth, M. A., & Gilbert, D. T. (2010). A wandering mind is an unhappy mind. *Science*, *330*, 932–932. https://doi.org/10.1126/science.1192439

Kvavilashvili, L., & Fisher, L. (2007). Is time-based prospective remembering mediated by self-initiated rehearsals? Role of incidental cues, ongoing activity, age, and motivation. *Journal of Experimental Psychology: General*, *136*, 112–132. https://doi.org/10.1037/0096-3445.136.1.112

Kvavilashvili, L., Kornbrot, D. E., Mash, V., Cockburn, J., & Milne, A. (2009). Differential effects of age on prospective and retrospective memory tasks in young, young-old, and old-old adults. *Memory*, *17*, 180–196. https://doi.org/10.1080/09658210802194366

Kvavilashvili, L., & Mandler, G. (2004). Out of one's mind: A study of involuntary semantic memories. *Cognitive Psychology*, *48*, 47–94. https://doi.org/10.1016/S0010-0285(03)00115-4

Mace, J. (2004). Involuntary autobiographical memories are highly dependent on abstract cuing: the Proustian view is incorrect. *Applied Cognitive Psychology*, *18*, 893–899. https://doi.org/10.1002/acp.1020

Mace, J., Bernas, R. S., & Clevinger, A. (2015). Individual differences in recognising involuntary autobiographical memories: Impact on the reporting of abstract cues. *Memory*, *23*, 445–452. https://doi.org/10.1080/09658211.2014.900083

MacKerron, G., & Mourato, S. (2013). Happiness is greater in natural environments. *Global Environmental Change*, *23*, 992–1000. https://doi.org/10.1016/j.gloenvcha.2013.03.010

Mazzoni, G., Vannucci, M., & Batool, I. (2014). Manipulating cues in involuntary autobiographical memory: Verbal cues are more effective than pictorial cues. *Memory & Cognition*, *42*, 1076–1085. https://doi.org/10.3758/s13421-014-0420-3

McDaniel, M. A., & Einstein, G. O. (2007). *Prospective memory: An overview and synthesis of an emerging field*. Thousand Oaks, CA: Sage Publications.

Miller, G. (2012). The smartphone psychology manifesto. *Perspectives on Psychological Science*, *7*, 221–237. https://doi.org/10.1177/1745691612441215

Monk, R. L., Heim, D., Qureshi, A., & Price, A. (2015). "I have no clue what i drunk last night" Using smartphone technology to compare in-vivo and retrospective self-reports of alcohol consumption. *PLOS ONE*, *10*, e0126209. https://doi.org/10.1371/journal.pone.0126209

Palmier-Claus, J., Ainsworth, J., Machin, M., Barrowclough, C., Dunn, G., Barkus, E., … Buchan, I. (2012). The feasibility and validity of ambulatory self-report of psychotic symptoms using a smartphone software application. *BMC Psychiatry*, *12*, 172.

Rasmussen, A. S., & Berntsen, D. (2011). The unpredictable past: Spontaneous autobiographical memories outnumber autobiographical memories retrieved strategically. *Consciousness and Cognition*, *20*, 1842–1846. https://doi.org/10.1016/j.concog.2011.07.010

Rasmussen, A. S., Ramsgaard, S. B., & Berntsen, D. (2015). Frequency and functions of involuntary and voluntary autobiographical memories across the day. *Psychology of Consciousness: Theory, Research, and Practice*, *2*, 185-205. https://doi.org/10.1037/cns0000042

Reason, J., & Mycielska, K. (1982). *Absent-minded?: The psychology of mental lapses and everyday errors*. Englewood Cliffs, NJ: Prentice Hall.

Schlagman, S., Kliegel, M., Schulz, J., & Kvavilashvili, L. (2009). Differential effects of age on involuntary and voluntary autobiographical memory. *Psychology and Aging*, *24*, 397–411. https://doi.org/10.1037/a0015785

Schlagman, S., & Kvavilashvili, L. (2008). Involuntary autobiographical memories in and outside the laboratory: How different are they from voluntary autobiographical memories? *Memory & Cognition*, *36*, 920–932. https://doi.org/10.3758/MC.36.5.920

Schwarz, N. (2012). Why researchers should think "real-time": A cognitive rationale. In M. R. Mehl & T. S. Conner (Eds.), *Handbook of Research Methods for Studying Daily Life* (pp. 22–42). The Guildford Press.

Shiffman, S., Stone, A. A., & Hufford, M. R. (2008). Ecological momentary assessment. *Annual Review of Clinical Psychology*, *4*, 1–32. https://doi.org/10.1146/annurev.clinpsy.3.022806.091415

Statt, N. (2016, April 18). Apple says the average iPhone is unlocked 80 times a day. Retrieved 24 April 2016, from http://www.theverge.com/2016/4/18/11454976/apple-iphone-use-data-unlock-stats

Stone, A. A., Shiffman, S., Schwartz, J. E., Broderick, J. E., & Hufford, M. R. (2003). Patient compliance with paper and electronic diaries. *Controlled Clinical Trials*, *24*, 182–199. https://doi.org/10.1016/S0197-2456(02)00320-3

Takarangi, M. K. T., Garry, M., & Loftus, E. F. (2006). Dear diary, is plastic better than paper? I can't remember: Comment on Green, Rafaeli, Bolger, Shrout, and Reis (2006). *Psychological Methods*, *11*, 119–122. https://doi.org/10.1037/1082-989X.11.1.119

Tennen, H., Affleck, G., Coyne, J. C., Larsen, R. J., & DeLongis, A. (2006). Paper and plastic in daily diary research: Comment on Green, Rafaeli, Bolger, Shrout, and Reis (2006). *Psychological Methods*, *11*, 112–118. https://doi.org/10.1037/1082-989X.11.1.112

Terry, W. S. (1988). Everyday forgetting: data from a diary study. *Psychological Reports*, *62*, 299–303. https://doi.org/10.2466/pr0.1988.62.1.299

Thai, S., & Page-Gould, E. (2017). Experiencesampler: An open-source scaffold for building smartphone apps for experience sampling. *Psychological Methods*. https://doi.org/10.1037/met0000151

Unsworth, N., Brewer, G. A., & Spillers, G. J. (2012). Variation in cognitive failures: An individual differences investigation of everyday attention and memory failures. *Journal of Memory and Language*, *67*, 1–16. https://doi.org/10.1016/j.jml.2011.12.005

Watkins, O. C., & Watkins, M. J. (1975). Buildup of proactive inhibition as a cue-overload effect. *Journal of Experimental Psychology: Human Learning & Memory*, *1*, 442–452. https://doi.org/10.1037/0278-7393.1.4.442

Whalen, C. K., Odgers, C. L., Reed, P. L., & Henker, B. (2011). Dissecting daily distress in mothers of children with ADHD: An electronic diary study. *Journal of Family Psychology*, *25*, 402–411. https://doi.org/10.1037/a0023473

Footnotes

[1] While loaned smartphones have many more features than PDAs, they are also similar to PDAs, since participants have to remember to carry them (in addition to their own mobile phone), or may need to receive training on how to use them (e.g., the elderly).

[2] The meaning of the term "smartphone" has evolved over the years. We define smartphones as Apple iPhone or Google Android-based phones, capable of running apps written by third-parties, and having Internet access and high-resolution touch-screens.

[3] One additional female smartphone participant did not return for two weeks, and the data on her phone indicated that most of her recordings were in the 14 days beyond the agreed 7-day period. Hence, her data were not included into the final sample of 60 participants reported here. This, however, illustrates a benefit of the smartphone diary.

[4] Two independent coders checked memory descriptions, and all were deemed to be autobiographical memories, probably due to very careful instructions and briefing.

[5] Just five of the 835 fully recorded memories were marked as "personal", one for each of two participants in the smartphone group, and one participant marking three as personal in a paper diary. Other items on the diary page for these "personal" memories were completed (e.g., ratings of concentration, vividness), and these entries were therefore included in the analyses.

[6] It is interesting that in 1-day Study 2, the paper-diary participants were as quick to record their memories as smartphone-diary participants. This improved compliance rate (at least for paper diaries) adds weight to the argument for shorter recording periods, where participants appear being more engaged with the study and record promptly.

Table 1

Mean Numbers (Standard Deviations) of Recorded and Acknowledged Involuntary Autobiographical Memories (IAMs) in Studies 1 and 2, and Everyday Memory Failures (EMFs) in Study 3, in Paper- and Smartphone-Diary Conditions.

	Condition	
	Paper-diary	Smartphone-diary
Study 1 (IAMs) – 7-day		
Fully recorded	18.03 (10.68)	9.52 (8.43)
Acknowledged	9.81 (9.86)	4.76 (4.84)
Study 2 (IAMs) – 1-day		
Fully recorded	7.54 (5.42)	3.87 (2.34)
Acknowledged	1.15 (1.89)	1.96 (2.06)
Study 3 (EMFs) – 7-day		
Fully recorded	10.60 (7.94)	5.63 (2.59)
Acknowledged	0.65 (1.17) [a]	1.06 (0.94) [b]

Note. [a] one outlier removed; [b] one outlier removed

Table 2

Mean Ratings (Standard Deviations) of Memory Characteristics as a Function of Condition (Paper- vs. Smartphone-Diary) in Study 1 and Study 2.

	Condition		
Study 1	Smartphone-diary (n = 29)	Paper-diary (n = 31)	$F(1, 58)$
Specificity [a]	0.78 (0.23)	0.68 (0.26)	2.56^{ns}
Vividness [b]	5.29 (0.93)	5.42 (0.59)	.47
Pleasantness now [c]	3.37 (0.78)	3.46 (0.43)	.27
Pleasantness then [c]	3.31 (0.83)	3.42 (0.50)	.43
Rehearsal [d]	2.65 (0.71)	2.41 (0.59)	2.10^{ns}
Study 2	Smartphone-diary (n = 23)	Paper-diary (n = 26)	$F(1, 47)$
Specificity [a]	0.73 (0.18)	0.80 (0.14)	2.27^{ns}
Vividness [b]	4.64 (1.49)	4.91 (0.97)	.60
Pleasantness now [c]	3.10 (0.99)	3.71 (0.59)	7.22**
Pleasantness then [c]	3.16 (0.88)	3.64 (0.71)	4.45*
Rehearsal [d]	2.41 (0.81)	2.78 (0.83)	2.44^{ns}

Note. [a] Involuntary memories were rated as specific or general. Means represent mean proportions of specific memories averaged across participants; [b] Ratings were made on a 7-point scale (1= *very vague*; 7 = *extremely vivid*); [c] Ratings were made on a 5-point scale (1 = *very unpleasant*; 2 = *quite unpleasant*; 3 = *neutral*; 4 = *quite pleasant*; 5 = *very pleasant*); [d] Ratings were made on a 5-point scale (1 = *never*, 2 = *once or twice*, 3 = *a few times*, 4 = *several times*, 5 = *many times*).
* $p = 0.04$; ** $p = 0.01$

Table 3

Mean Proportions (Standard Deviations) of Memories with Reported Internal, External and No Triggers by Condition (Paper- vs. Smartphone-Diary) in Study 1 and Study 2.

	Trigger Type		
	Internal	External	No Trigger
Study 1			
Smartphone-diary	0.18 (0.17)	0.68 (0.21)	0.14 (0.20)
Paper-diary	0.16 (0.13)	0.74 (0.17)	0.10 (0.10)
Study 2			
Smartphone-diary	0.10 (0.13)	0.83 (0.19)	0.07 (0.14)
Paper-diary	0.13 (0.15)	0.75 (0.24)	0.12 (0.19)

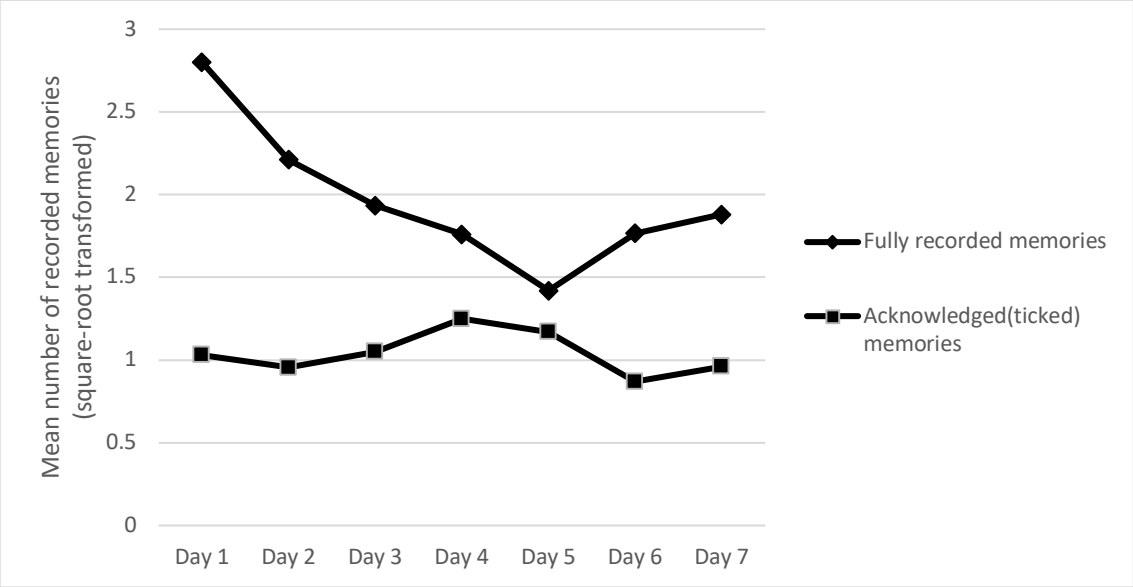

Figure 1. *Mean numbers of fully recorded and acknowledged memories (square-root transformed) pooled across paper- and smartphone diary conditions in Study 1.*

Appendix 1

Example paper-diary page, used in Study 1 on involuntary autobiographical memories.

1. a) When did you have a memory? Date_____Time_____AM/PM

1. b) When did you record it? Date_____Time_____AM/PM

2. Describe your memory. What was it about?

3. How vivid is your memory?

❍	❍	❍	❍	❍	❍	❍
Very vague						extremely vivid

4. Was the memory triggered by something

❍ in your thoughts ❍ in your environment ❍ there was no trigger

If a trigger, what was it?

Please estimate the time between the trigger and the memory:

5. What were you doing?

6. How much were you concentrating on this activity?

❍	❍	❍	❍	❍
Not at all				Fully concentrating

7. How pleasant or unpleasant is the memory now?

❍	❍	❍	❍	❍
Very unpleasant	Quite unpleasant	Neutral	Quite pleasant	Very pleasant

8. How pleasant was the original event?

❍	❍	❍	❍	❍
Very unpleasant	Quite unpleasant	Neutral	Quite pleasant	Very pleasant

9. Is the memory of a general or specific event?

❍ General ❍ Specific

10. When did the original event occur?

11. Have you ever had this memory before?

❍	❍	❍	❍	❍
Never	once or twice	a few times	several times	many times

Appendix 2

Example paper-diary page, used in Study 3 on everyday memory failures.

When did you have a memory error? Date: _____ Time: _____ AM/PM
Or when did you realise you made an error?

When did you record it here? Date: _____ Time: _____ AM/PM

Describe your memory error (what it was, what doing, where you were):

What was your mood immediately before the error? (tick)

☐	☐	☐	☐	☐	☐
Very unhappy		Neutral		Very happy	Don't Know

How relaxed or stressed were you immediately before the error? (tick)

☐	☐	☐	☐	☐	☐
Very relaxed		Neutral		Very stressed	Don't Know

How serious was the memory lapse? (tick)

☐	☐	☐	☐	☐
Insignificant	Minor	Somewhat significant	Significant	Very significant / potentially dangerous

Were there or could there have been any consequences?

How upset are you by the memory lapse?

☐	☐	☐	☐	☐
Not at all upset	A little	Somewhat	Quite	Very upset

Describe the emotions you felt in response to your lapse, if any:

If you recovered from the error, describe when and how: