Memory for staged events: supporting older and younger adults' memory with SenseCam

Ali Mair, Marie Poirier, and Martin A. Conway

Department of Psychology, City, University of London, UK

Address for correspondence:

Ali Mair

Department of Psychology and Sports Sciences

University of Hertfordshire

College Lane

Hatfield, AL10 9AB

Tel. +44 (0) 20 7040 4597

Email: a.mair2@herts.ac.uk

Abstract

Two experiments measured the effect of retrieval support provided by a wearable camera, SenseCam, on older and younger adults' memory for a recently experienced complex staged event. In each experiment participants completed a series of tasks in groups and the events were recalled two weeks later, after viewing SenseCam images (experimental condition) or thinking about the event (control condition). When IQ and education were matched, young adults recalled more event details than older adults, demonstrating an age-related deficit for novel autobiographical material. Reviewing SenseCam images increased the number of details recalled by older and younger adults, and the effect was similar for both groups. These results suggest that memory can be supported by the use of SenseCam, but the age-related deficit is not eliminated.

KEYWORDS: Ageing; autobiographical memory; episodic memory; lifelogging; wearable technology.

Memory for staged events: supporting older and younger adults' memory with SenseCam

The present study examines the effect of retrieval support, in the form of images captured by a wearable camera (SenseCam), on older and younger adults' memory for complex staged events. Previous work has shown promise for the use of wearable cameras as a prosthesis for memory impairment in people with amnesia (Berry, Kapur, Williams et al., 2007; Loveday & Conway, 2011; Pauly-Takacs, Moulin, & Estlin, 2011), mild cognitive impairment (Browne, Berry, Kapur, et al., 2011), and Alzheimer's disease (E. Woodberry, Browne, Hodges, Watson, Kapur, & K. Woodberry, 2015). In addition, a number of studies have shown that images captured by wearable cameras can be used to cue memory in healthy young adults (Finley, Brewer & Benjamin, 2011; Kalnikaité, Sellen, Whittaker and Kirk, 2010; Mair, Poirier, & Conway, 2017; Sellen, Fogg, Aitken, et al., 2007; but see Seamon, Moskowitz, Swan, et al., 2014). One group that stands to benefit considerably from this type of technology is older adults, for whom episodic memory impairments are well-established; yet little work so far has addressed this potential. In the present paper, we therefore extend the use of wearable camera technology to healthy older adults in a controlled study designed to imitate a complex everyday event.

Age-related memory impairments. Over the last 50 years, a wealth of literature has demonstrated age-related declines in performance on a range of episodic memory tasks. Of relevance to the present study is the relative impairment older adults exhibit in memory for details of specific, personally-experienced events (Kvavilashvili, Mirani, Schlagman, Erskine & Kornbrot, 2010; Levine, Svoboda, Hay, Winocur & Moscovitch, 2002; Mueller-Johnson & Ceci, 2004; Piolino, Coste, Martinelli et al., 2010; Piolino, Desgranges, Benali & Eustache, 2002; Piolino, Desgranges, Clarys, et al., 2006; West & Stone, 2013). Typically, older adults recall fewer episodic or event-specific details than younger adults, but retain access to more generic event information such as scripts and schemas (Light & Anderson, 1983; Rosen, Caplan, Sheesley, Rodriguez, & Grafman, 2003). In some cases, prior knowledge may be able to compensate for age-related declines in episodic memory performance (Badham & Maylor, 2015; Badham, Hay, Foxon, Kaur, & Maylor, 2016; Umanath & Marsh, 2014), however when prior knowledge about a particular type of event is limited (e.g. if the event type is novel), older adults may be more likely to need external memory support.

Previous work has investigated the effect of reviewing photographs on participants' memory for complex events. In one such study, photograph review increased the amount older and younger adults recalled about videos seen two days earlier, but the benefit for older adults was smaller than for younger adults if images were only reviewed once; when images were reviewed three times, older and younger adults benefitted equally (Koutstaal, Schacter, Johnson, Angell, & Gross, 1998). In contrast, older adults benefitted more from photograph review than younger adults when the task was to recall actions they had personally performed two days earlier, but the recall benefit for *details* about those actions was equal for older and younger adults (Koutstaal, Schacter, Johnson, & Galluccio, 1999).

Wearable cameras as memory support. The use of wearable cameras to support event memory is a relatively recent enterprise, yet already there are numerous examples of the potential for such technology to increase the amount of event information that individuals can recall (Silva, Pinho, Macedo & Moulin, 2016). In the present paper we are concerned with SenseCam (SC) technology, which has been the most widely-used to date. SC is worn around the neck and captures still images automatically, from the wearer's perspective, in response to changes in external stimuli such as light, motion, and acceleration. In practice, the device captures one image approximately every 9-10 seconds, and these images can be uploaded to a computer and reviewed in sequence, providing a richly detailed and objective record of a previously experienced event (Hodges, Berry & Wood, 2011). The images captured by SC are wide-angled to maximise the field of view (which gives the images a "fish-eye" quality), and relatively low resolution to facilitate storage of up to 30,000 files. Several features of SC images differentiate them from photographs captured by typical use of regular digital cameras. For example, SC generates a much larger number of images, and images are passively captured, sequentially ordered, temporally compressed, capture a wide field of view, and are triggered in response to potentially important environmental changes. Hodges et al. (2011) suggested that these features are compatible with normal memory, and therefore may render SC particularly effective for memory support.

Despite the success of SC in supporting recall in people with memory impairments due to neurological conditions (Berry et al., 2007; Browne et al., 2011; Loveday & Conway, 2011; Pauly-Takacs et al., 2011; Woodberry et al., 2015), there is as yet little work investigating the effect of SC on memory in healthy older adults. One study found that reviewing personal SC images from three previous days improved older and younger adults' performance on a variety of standardised cognitive tests, including an unrelated autobiographical memory task (Silva, Pinho, Macedo and Moulin, 2013). This suggests that SC may provide a general benefit to cognitive function, however in that study memory for the reviewed events was not tested. In another study, participants reviewed SC-like images of a previously visited museum, although SC was not worn during the museum tour itself. On a subsequent task examining recognition of the museum exhibits, both older and younger adults scored more hits, but also more false alarms, when the perspective and temporal order of the reviewed images was matched more closely to the experience at encoding (St Jacques, Montgomery, & Schacter, 2015).

To our knowledge, only one previous study has measured the effect of SC on older adults' recall of the SC-reviewed events (Mair et al., 2017). In that study, participants self-selected typical events from everyday life, which were prospectively sampled using SC. Two weeks later, the sequence of SC images was presented as a retrieval cue, and participants were instructed to describe everything they could remember about the event. The results showed that in both older and younger adults, SC review increased the number of recalled event details relative to an unsupported control condition. However, there was no difference between older and younger groups' event memory in the control condition, and the effect of SC was the same for both groups. It is therefore not yet clear whether SC can compensate for age-related memory impairment. In the present paper we investigate whether SC can improve event memory in older adults in a task in which there is likely to be a baseline impairment: recall of a novel staged event.

Experiment 1

In Experiment 1, we aimed to measure the effect of reviewing SC images on the number of details older and younger adults could recall about the reviewed event. An additional aim of Experiment 1 was to investigate the effect of the temporal order of images on subsequent recall. As yet, little is known about the determinants of SC's success. Hodges et al. (2011) proposed two competing hypotheses: first, they suggested that the sheer amount of information available in a sequence of SC images makes it likely that at least some of the images will reinstate something that was encoded in memory. Secondly, Hodges et al. (2011) suggested that the mode of capture and review of SC images is particularly compatible with human memory. That is, SC images are visual, passively captured, time-compressed, sequentially ordered, from the wearer's perspective, and so on. This idea is not inconsistent with the generally-accepted view that human memory is reconstructive (Bartlett, 1932; Conway & Pleydell-Pearce, 2000; Rubin & Umanath, 2015); it does not suggest that the mode of SC's operation is identical to that of human memory, but rather that certain features of the two systems overlap. Some support for the compatibility hypothesis comes from healthy young adults, for whom passive versus active photo capture, own versus other perspective during review, and forward versus random temporal order of image sequences has been shown to lead to better memory performance (Sellen et al., 2007; St Jacques & Schacter, 2013). In addition, Mair et al. (2017) found a small but significant effect of temporal order, such that older and younger adults recalled fewer event details if they reviewed SC images of an event in random order compared to forward order. However, two of the above studies did not measure memory for rich event-specific detail (Sellen et al., 2007; St Jacques & Schacter, 2013), while in the third participants recalled the event at the same time as reviewing SC images (Mair et al., 2017), so that the order of the recall would also have been random. This poses a potential problem for interpretation, since recall in forward order produces more details than recall in an alternative order (Anderson & Conway, 1993). In Experiment 1, we therefore compared the effect of forward order and random order image sequences in a procedure in which recall took place after image review. We compared these two review conditions to a control condition in which participants did not review any SC images. Based on previous work, we predicted that both older and younger adults would recall more event details after reviewing SC images compared to the control condition. Predictions about the effect of temporal order were derived from the two competing hypotheses proposed by Hodges et al. (2011). If the compatibility of SC with human memory is important, then reviewing the images in random order should lead to worse subsequent recall performance than reviewing the images in forward order. On the other hand, if the important factor is the sheer amount of information present in the cue, then random and forward order review should lead to comparable recall performance.

Method

Participants. Eighteen young adults (age 19-32; *M*=23.72, *SD*=3.91; 15 female) and 25 older adults (age 64-83; *M*=72.32, *SD*=5.68; 19 female) participated for a payment of £8 per hour. Young adults were recruited via posters displayed around the university buildings and via social media. Twelve of the young adults were students at City University London, and the remaining six were external participants. Older adults were recruited from a pool of participants who had previously responded to a local newspaper advertisement and expressed an interest in taking part in memory experiments. All participants were native English speakers, with self-reported normal or corrected-to-normal vision. Older adults were screened for dementia using the Mini Mental State Examination (MMSE) using a cut-off of 24 (Folstein, Folstein & McHugh, 1975). The mean MMSE score was 28.38 (SD=1.50; range 25-30; the participant who scored 25 refused to attempt the subtraction question worth 5 marks).All participants completed the Geriatric Depression Scale (GDS; Yesavage, Brink, Rose, et al., 1983), which provided a rough estimate of depressive symptoms that may reduce memory specificity (Birch & Davidson, 2007). The National Adult Reading Test (NART; Nelson, 1982) was administered to all participants to give an estimate of IQ (Bright, Jaldow & Kopelman, 2002), and the number of years of formal education each participant had received was also recorded. The NART was used on this occasion because of the short amount of time taken to administer the instrument, which was necessary because participants were tested in groups. IQ was calculated from the number of reading errors using the following formula: Full Scale IQ = 128 - 0.83 x NART error score (Nelson, 1982). Younger adults had more years of education (M=15.83, SD=2.57) than older adults (M=13.67, SD=3.21; t(40)=2.35, p=.02) but older adults had higher IQ scores (M_{older}=116.10, SD=7.56; $M_{younger}$=109.82, SD=6.93; t(40)=2.73, p=.009). There was no difference in GDS score between groups (M_{older}=8.24, SD=6.24; $M_{younger}$=6.56, SD=3.09; t(41)=1.05, p=.25).

Design. The design of Experiment 1 involved two rounds of measurement, both in the form of written questionnaires (see *Appendix A*). In the first, a baseline measure of memory performance was obtained by recording the number of details that were recalled about the whole event, prior to any experimental manipulation. The second measure was related to the

SC manipulation: a 2 (age group: young vs. old) x 3 (SC condition: control vs. random temporal order vs. forward temporal order) mixed factorial design was employed; in effect, the event to be recalled was split into three sections in order to test retrieval condition under repeated measures. The dependent variable in the second round of measurement was the number of *additional* details that were recalled (i.e. new details that were not recalled at baseline). The control condition measured the number of additional details recalled in the second questionnaire without viewing SC images, while the random order and forward order conditions measured the number of details freely recalled after reviewing photos from the event in random order and forward order, respectively. The number of episodic and semantic details was recorded, as well as incorrect details and source memory errors.

Materials and Procedure

Encoding session. The recording event for the participants consisted of a staged event, in which small groups of participants visited three separate rooms, and took part in a series of group activities in each room. There were six separate encoding sessions, each attended by up to nine participants (min. =6). At each session, participants were split into three groups of 2 or 3, and each participant was provided with a SC which they wore for the duration. Allocation to groups was opportunistic, based on where participants chose to sit in the waiting area upon arrival, and consequently the natural social bonds that had begun to form. There were 18 groups in total: 7 mixed-age groups, 7 older adult groups, and 4 younger adult groups. The event was designed to standardise the material that each participant would be asked to remember at test, as well as to ensure that the experience that participants would be reviewing in each of the review conditions would be comparable. For the latter reason, the event was split into three sections, each of which took place in a separate room with its own set of three tasks.

There were three types of task in each room: one visual task (name the flag), one auditory (general knowledge questions), and one problem-solving task that was different in each room. The flags task involved the presentation of coloured flags on an overhead projector. Each flag belonged to a different country, and participants were asked to name the countries the flags belonged to. In the general knowledge task, the experimenter read the questions aloud and asked participants to guess the answer. The problem-solving tasks were not question-and-answer based, and instead involved participants working together to achieve a particular goal. The tasks are described in more detail in *Appendix B*. The rooms were broadly colour-themed (red, blue and green) to minimise confusion at test because of the similarity of tasks in each room, and the theme was reinforced through coloured pictures on the walls of each room, as well as the colour of the flags in the flags tasks. Each room was managed by a task leader, who wore clothing to match the colour of their room.

Tasks were designed to be difficult so that recall advantages associated with prior knowledge of the material were not conferred on either group, and the materials for the tasks were determined in a pilot questionnaire which was completed online by 37 participants of mixed ages. In both general knowledge and flags tasks, participants were asked to provide an individual answer or guess first and then subsequently to choose an answer to submit as a group. This manipulation was designed to ensure that all participants attended to the stimuli and any failure to remember particular details at test was not due to attentional differences during encoding. The request for a group answer was included to encourage interaction between the participants in each group, thereby increasing the social autobiographical aspect of the task, as well as providing additional material that could subsequently be recalled at test. After group answers were provided, the task leader told each group some facts about the question they had just answered. This was designed to increase the volume of material that could be recalled at test, and the material was scripted to ensure that the same material was presented to each group in the same way. If the group gave an incorrect answer to any question, the correct answer was provided by the task leader.

Each group was allocated to one room (red, blue or green) to start. The groups moved around each of the three rooms, but the task leaders always remained with their own room. The tasks within each room were administered in random order, and each group spent 15-20 minutes completing the activities before moving to the next room. The order of the rooms was counterbalanced across sessions. Each room was video-recorded to enable participants' memories of non-scripted details to be checked for accuracy.

Recall sessions. Participants returned for the recall session fourteen days after their recording session, with the exception of one participant in the young adult group who was unable to attend the group recall session and came in one day earlier. The recall session was held in a computer lab, and participants were met and debriefed individually. The recall test was a semi-structured written questionnaire, divided into one section per room, which asked participants to remember details about the event. Most questionnaires were administered on the computer, although older participants who were not comfortable using computers were provided with a paper copy (n=9). Participants were reminded of the colour theme and the location of each room relative to the others, as well as the name of the task leader for each room. Questions probed for details about the visual environment, the task leader and the tasks in each room (see *Appendix A*), and participants were asked to write in as much detail as possible. Participants first filled out the full questionnaire for all three rooms (baseline; T1), in which the sections were ordered to match the order in which participants had visited the rooms two weeks earlier. Participants' personal SC photos captured within each room were then reviewed in one of three experimental conditions: control, random order, or forward order. In random and forward order conditions, participants reviewed all of the images from the corresponding room in random temporal order and forward temporal order, respectively.

Image review was self-paced, with participants pressing a key on the keyboard to advance through the image sequence. Participants were not provided with any additional instructions for review. In the control condition, participants did not review any images, and instead were asked to spend a few minutes thinking about the room in question before moving onto the T2 questionnaire. Immediately after the review of each room participants filled out the corresponding section of the questionnaire again on a new blank copy (T2 recall), and were instructed to write down all remembered details, including new details and those that had already been reorted at T1.

Coding strategy. Memories at both T1 and T2 were coded for four types of information: episodic and semantic details, source errors, and incorrectly recalled details. A master response sheet was created to ensure the scripted details were coded consistently, and the total number of details in each category was tallied. Episodic details were those that referred to event-specific information, including objects that were present, physical descriptions of the rooms and task leaders, actions (e.g. moving three matchsticks to create a particular design), interactions with others, internal thoughts and feelings, and more abstract facts or concepts encountered during the event (e.g. remembering that the small pieces of paper punched out by a hole-puncher are called chads). Episodic details were further marked as visual (e.g. objects, physical descriptions, etc.) or non-visual (e.g. actions, interaction with other participants, internal thoughts, etc.) for a secondary analysis. It should be noted that, where possible, the episodic category contained details that were verifiable by checking against scripts and video recordings. However, participants were encouraged to also recall thoughts and feelings that they had at the time, which we were unable to verify. Semantic details were those that were not linked to a specific time and place (e.g. *my general* *knowledge is poor*[1]). Source errors were details that were correct in content, but recalled as part of the wrong room, while incorrect details were those that either contained errors (e.g. recalling ten letters in the word wheel rather than nine) or those that referenced something that was not present or did not happen. The threshold for incorrect responses necessarily varied with the type of information (e.g. the circumference of the earth in miles could be rounded to the nearest thousand, but the colours and shapes on a flag had to be recalled correctly); acceptable answers were predefined in the master coding sheet to ensure consistency across the sample. See *Appendix C* for a coded example of a participant's response. For T2 questionnaires details were classed as new if they had either not been previously reported, or had been reported differently on the T1 questionnaire (e.g. an incorrect detail that was corrected after reviewing SC images). In order to measure the reliability of the coding strategy, three additional raters coded a subset of six participants' questionnaires (three older adults and three younger adults). A two-way random intraclass correlation was calculated for the four coders' responses (r_{ic}=.99; 95% CI = .96, 1.00), which showed that agreement was good.

Results

Preliminary analysis. Semantic details were excluded from further analysis because only 7 participants reported any. Partial correlations controlling for age found that both IQ (r=.45, p=.004) and education (r=.40, p=.01) were positively correlated with the number of episodic details recalled at T1, but were not correlated with source errors or incorrect details (all *p*s>.65). In order to control for any differences in performance that might be attributable to differences in IQ or education, we balanced our sample on these measures[2]. Five older

---

[1] Note that this is classed as a semantic detail because it was reported as a comment during the recall test, and not as something that the participant recalled thinking at the time of encoding two weeks earlier.

[2] We chose this approach, rather than entering IQ into the analyses as a covariate, because IQ was statistically different between groups. Since participants were not randomly allocated to groups, this difference cannot be thought of as random, and an analysis of covariance under such circumstances would be invalid (G. A. Miller &

adults and two young adults were excluded from the sample on the basis of their IQ and education scores. The excluded older adults were those who had IQ scores above the group average but education scores below the group average, and the excluded young adults had the opposite pattern of results, with IQ scores below the group average but education scores above the group average. We also excluded one young adult and one older adult for whom IQ or education data were missing. T-tests showed that in the matched sample (n=34; 15 younger, 19 older) there was no group difference in IQ ($M_{younger}$=110.13, SD=7.34; M_{older}=114.97, SD=7.38; $t(32)$=1.90, p=.066) or education ($M_{younger}$=15.33, SD=2.50; M_{older}=13.95, SD=3.50; $t(32)$=1.29, p=.205).

T1 data. Data from the T1 questionnaires were averaged across rooms. The mean number of episodic details, incorrect details, and source errors is presented in *Table 1*. Since incorrect details and source errors were near floor, they were analysed separately from episodic details. We therefore compared older and younger adults' episodic recall in an independent groups t-test, which showed that younger adults recalled significantly more details than older adults ($t(32)$=2.96, p=.006). Memory errors were then entered into a 2 (error type: source error vs. incorrect detail) x 2 (age group) ANOVA, which showed no significant effect of age ($F(1,32)$=.87, p=.36, η_p^2=.03, MSE=3.95) or error type ($F(1,32)$=1.15, p=.29, η_p^2=.04, MSE=2.79), and no significant interaction between age group and error type ($F(1,32)$=3.41, p=.07, η_p^2=.10, MSE=2.79).[3] However, we also looked at the

---

Chapman, 2001). Using covariates in this way is liable to distort the effect that is being measured, since there is an overlap in the variance explained by the group and the covariate (Cochrane, 1957). For this reason, we excluded participants as described above, however for clarity the results for the full sample are presented in footnotes for each analysis.

[3] The pattern of results at T1 was similar for the full sample, with a significant effect of age on episodic recall ($t(41)$=3.71, p=.001), but not memory errors ($F(1,41)$=2.00, p=.17, η_p^2=.05, MSE=4.84). In the full sample there was no main effect of error type ($F(1,41)$=1.97, p=.17, η_p^2=.05, MSE=2.86), but age group interacted with error type ($F(1,41)$=4.52, p=.04, η_p^2=.10, MSE=2.86). Post-hoc t-tests corrected for multiple comparisons did not reach significance for either error type. There was, however, a significant effect of age on the proportion of memory errors ($F(1,41)$=8.31, p=.006, η_p^2=.17, MSE=.01), similar to what was observed in the matched sample.

proportion of errors-to-total recall, in order to control for differences in generativity (see *Table 1*). These data were entered into a 2 (age group) x 2 (proportion of error type) ANOVA. Proportionally, older adults made significantly more recall errors than younger adults at T1 ($F(1,32)=4.86$, $p=.04$, $\eta_p^2=.13$, $MSE=.01$; $M=.13$, $SD=.08$ vs. $M=.07$, $SD=.08$). There was no significant interaction between age group and error type ($F(1,32)=.91$, $p=.35$, $\eta_p^2=.03$, $MSE=.008$).

[Table 1 about here]

T2 data. We next looked at the number of new details produced at T2 (see *Table 2*). No new source errors or semantic details were added, therefore the following analyses are presented for episodic and incorrect details only. Detail types were again analysed separately due to the number of incorrect details approaching floor. New episodic details were analysed in a 2 (age group) x 3 (condition: control vs. random vs. forward) ANOVA, which found no main effect of age ($F(1,31)=.09$, $p=.77$, $\eta_p^2<.01$, $MSE=36.38$), and no age by condition interaction ($F(2,62)=1.89$, $p=.16$, $\eta_p^2=.06$, $MSE=15.13$). There was, however, a main effect of condition ($F(2,62)=12.03$, $p<.0005$, $\eta_p^2=.28$, $MSE=15.13$), which post-hoc pairwise comparisons with Bonferroni corrections revealed to be driven by significant differences between forward order ($M=7.63$, $SD=4.80$) and control ($M=3.53$, $SD=3.82$; $p<.0005$), and random order ($M=7.66$, $SD=5.52$) and control ($p=.002$). There was no difference between forward order and random order review conditions ($p=1.00$).[4] A similar analysis was carried out on T2 incorrect details. There was no main effect of age ($F(1,31)=.86$, $p=.36$, $\eta_p^2=.03$,

[4] In the full sample the T2 results for episodic details were similar. There was a main effect of condition ($F(2,82)=18.76$, $p<.0005$, $\eta_p^2=.31$, $MSE=13.79$), but no main effect of age ($F(1,41)=.42$, $p=.52$, $\eta_p^2=.01$, $MSE=54.06$), and no interaction ($F(2,82)=2.21$, $p=.12$, $\eta_p^2=.05$, $MSE=13.79$).

MSE=2.05) or condition (*F*(2,62)=1.17, *p*=.32, η_p^2=.04, *MSE*=1.36), and no interaction (*F*(2,62), *p*=.89, η_p^2<.01, *MSE*=1.36).[5]

[Table 2 about here]

We also analysed the number of memory errors as a proportion of total T2 recall (see *Table 2*). Proportions could not be calculated in at least one condition for a total of six participants (three younger and three older), because they failed to add any new details at T2. These six participants were excluded from all three conditions and the following analysis therefore involves 12 younger adults and 16 older adults. A 2 (age group) x 3 (condition) found no significant effects of age (*F*(1,26)=2.41, *p*=.13, η_p^2=.09, *MSE*=.06) or condition (*F*(2,52)=.10, *p*=.90, η_p^2<.01, *MSE*=.04), and no age by condition interaction (*F*(2,52)=.31, *p*=.73, η_p^2=.01, *MSE*=.04).[6]

These results suggest that SC increases the number of episodic details both young and older participants recall about an event, but does not affect the number of memory errors. One possible argument, however, is that in the T2 questionnaires participants were reporting what they had just seen in the pictures rather than what they remembered of the original event. Since it was not possible to determine the source of the remembered details, a second round of analysis was carried out on the episodic details after excluding all of the visual details. The nonvisual data are presented in *Table 2*. This yielded a conservative measure of the number of original details recalled, since it is highly probable that at least a proportion of visual details were not depicted in the images themselves. Analysis of nonvisual details in a 2 (age) x 3 (condition) ANOVA showed a small but significant main effect of condition

[5] In the full sample, there was no effect of age (*F*(1,41)=.64, *p*=.43, η_p^2=.02, *MSE*=2.14) or condition (*F*(2,82)=1.28, *p*=.28, η_p^2=.03, *MSE*=1.94) on the number of T2 incorrect details, and no age by condition interaction (*F*(2,82)=.08, *p*=.92, η_p^2<.01, *MSE*=1.94).

[6] The results of the proportional analysis in the full sample were also similar to the IQ-matched sample; there were no effects of condition (*F*(2,64)=.17, *p*=.84, η_p^2<.01, *MSE*=.03) or age (*F*(1,32)=3.22, *p*=.08, η_p^2=.09, *MSE*=.05) and no interaction (*F*(2,64)=.74, *p*=.48, η_p^2=.02, *MSE*=.03).

(F(2,62)=3.21, p=.047, η_p^2=.09, MSE=5.63), but pairwise comparisons revealed no reliable differences between control, random order and forward order conditions (all ps>.05). There was no main effect of age (F(1,31)=2.54, p=.12, η_p^2=.08, MSE=11.01) and no interaction between age and condition (F(2,62)=.73, p=.49, η_p^2=.02, MSE=5.63).[7]

Discussion

The results of Experiment 1 showed that, at baseline, younger adults' event memory was almost twice as detailed as older adults' memory of the same event. This age-related deficit in event memory is consistent with previous studies (Kvavilashvili et al., 2010; Levine et al., 2002; Mueller-Johnson & Ceci, 2004; Piolino et al., 2010; Piolino et al., 2002; Piolino et al., 2006; West & Stone, 2013). In addition, both older and younger adults' event memory benefitted from the use of SC as a retrieval cue, and the effect of SC was similar for both groups. While this effect was consistent with our previous findings (Mair et al., 2017), it suggests that older adults do not benefit disproportionately from SC review. That is, SC supported older adults' memory, but did not eliminatethe age-related deficit.

Temporal order. In Experiment 1, we investigated the effect of randomising the SC image sequence, to test the hypothesis that SC is an effective memory aid because it is particularly compatible with human memory (Hodges et al., 2011). According to this compatibility hypothesis, one such way in which SC images resemble natural memory is their organisation in temporal order. In the present experiment, both forward- and random-order images supported recall equally well, therefore we did not find any support for this hypothesis. This is in contrast to previous findings, which have shown a detrimental effect of

[7] In the full sample similar results were obtained for the analysis of nonvisual details. There was a main effect of condition (F(2,82)=6.97, p=.002, η_p^2=.15, MSE=5.47), but no effect of age (F(1,41)=1.76, p=.19, η_p^2=.04, MSE=16.74) and no age by condition interaction (F(2,82)=1.13, p=.33, η_p^2=.03, MSE=5.47). Pairwise comparisons found significantly higher nonvisual recall in random (M=3.45, SD=3.28, p=.02) and forward (M=3.78, SD=3.30, p=.007) conditions compared to control (M=1.99, SD=2.61), but no difference between random and forward conditions (p=1.00).

random-order image review compared to forward-order review (Mair et al., 2017; St Jacques et al., 2015; St Jacques & Schacter, 2013). One possibility is that temporal order is important only when there is less information available in the cue, such as in the studies by St Jacques and colleagues, where only a subset of the total number of captured images were presented to participants. When all images are reviewed, it may be that the large amount of information in the cue is sufficient to cue memory regardless of the order of presentation, as suggested by the alternative hypothesis proposed by Hodges et al. (2011). On the other hand, in our previous study (Mair et al., 2017), we compared forward- and random-order presentation for the review of full sets of images (i.e. a large amount of information in the retrieval cue) and found a small but significant detriment of random order presentation. There are at least two possible explanations for the discrepancy between our previous and present findings concerning temporal order. Firstly, it is possible that the temporal order effect in Mair et al. (2017) was attributable to disruption of recall order in the random condition (Anderson & Conway, 1993), since participants recalled events at the same time as reviewing the images. Alternatively, it may be that the present experiment lacked sufficient power to detect a true, small effect of temporal order, particularly since the sample size was reduced in order to match IQ between groups. In any case, the results presented both here and in our previous study (Mair et al., 2017) suggest that, when participants review full image sets, any effect of temporal order is minimal compared to the general recall benefit provided by SC.

Recall of “off-camera” information. Silva et al. (2016) proposed that an important research question is whether wearable cameras are able to cue recollection of information that is not depicted in the images. In Experiment 1, we attempted to address this by excluding visual details from the analysis of post-review recall. This may have underestimated event memory for two reasons: firstly, because visual details are a major component of autobiographical memory (M. A. Conway & Pleydell-Pearce, 2000; Galton, 1879). Secondly, most of the tasks participants were asked to complete had a strong visual element, and description of these tasks would have encouraged the recall of visual information. One of the questions included in the questionnaire asked explicitly that participants described the visual environment of the room in question, and two others ("Describe the task leader" and "Do you remember any of the flags from the [red/blue/green] room?") carried an implication that recall of visual information would be appropriate. Consequently, exclusion of the visual details considerably reduced the amount of information recalled across all conditions.

Nonetheless, there was a small effect of review condition on nonvisual details, suggesting some recall of "off-camera" information, although no differences were detected between individual conditions. We suggest that the visual nature of the tasks in Experiment 1 contributed to the visual focus of the recall, and that to address the concern of Silva et al. (2016) that wearable cameras should be able to cue "something more" than what is shown in the images, it is necessary to measure memory for tasks with more nonvisual elements.

Experiment 2

In Experiment 2 we again aimed to test the effect of SC review on older and younger adults' event memory, this time with a focus on the recall of "off-camera" information. To reflect this updated focus, three key alterations were made to the study procedure. Firstly, the tasks completed by participants during the event were changed to increase the amount of nonvisual information available to recall. The tasks in Experiment 2 also departed from the question-and-answer format of Experiment 1, and participants were encouraged to interact more with the task materials and with each other. The second change was that participants took part in larger groups of up to 10 individuals, compared to groups of two or three individuals in Experiment 1. This change was intended to encourage interaction and conversation, which should also increase the capacity for nonvisual recall at test. Thirdly, in

Experiment 2 we measured recall in a one-to-one interview rather than a written questionnaire, in which we could ensure participants fully understood the review and recall instructions. As in Experiment 1, we predicted that SC review would lead to better event recall in older and younger adults. Additionally, we predicted that SC review would lead to greater recall of nonvisual details than unsupported review, in both younger and older adults. In Experiment 2, we did not pursue the random order condition further, since Experiment 1 and our previous work (Mair et al., 2017) suggested at best only a small effect of temporal order is observed when all images were reviewed.

Method

Participants. Seventeen young adults (age 18-32; *M*=24.29, *SD*=4.70; 14 female) and 19 older adults (age 66-85; *M*=71.00, *SD*=4.18; 15 female) participated in return for a payment of £20. Young adults were predominantly undergraduate students recruited via an online sign-up system, while older adults were recruited from a panel of individuals who had previously responded to an advertisement in a local newspaper. Two of the young adults and six of the older adults also participated in Experiment 1.All participants had self-reported normal or corrected-to-normal vision. As in Experiment 1, older participants were screened using the MMSE and all participants scored well above the cut-off point of 24 suggested by Folstein et al. (1975)(M=28.89, SD=.88, range=27-30). Both groups completed the Geriatric Depression Scale (GDS; Yesavage et al., 1983) as an indicator of depressive symptoms that may be associated with reduced memory specificity. The 2 subscale version of the Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler, 1999) was administered to both groups to give an estimate of IQ[8].

[8] The NART was used to measure IQ in Experiment 1 because of the time restriction caused by testing participants in groups; however our preferred instrument is the WASI, which was used in Experiment 2. We did not collect both NART and WASI scores for either of the samples described here, however both were measured

Years of education did not differ significantly between younger (*M*=16.88, *SD*=3.10) and older adults (*M*=15.74, *SD*=3.35; $t(34)=1.06$, $p=.29$), but older adults (*M*=125.00, *SD*=8.89) had significantly higher IQs than younger adults (*M*=109.24, *SD*=13.74; $t(34)=4.13$, $p<.0005$). Young adults (*M*=8.76, *SD*=6.58) also showed significantly more indicators of depressive symptoms than older adults (*M*=4.37, *SD*=3.82; $t(34)=2.41$, $p=.02$). Recall of episodic details at baseline was significantly correlated with IQ when controlling for age group ($r=.54$, $p=.001$), but there was no significant correlation between correct recall and education ($r=.20$, $p=.26$) or GDS ($r=-.23$, $p=.18$). As in Experiment 1, to control for group differences in IQ we balanced the sample by excluding the five young adults with the lowest IQ scores and the eight older adults with the highest IQ scores. This adjustment ($n=24$; 12 younger) left no significant group differences in IQ ($t(22)=1.15$, $p=.26$), education ($t(22)=1.83$, $p=.08$) or GDS ($t(22)=.22$, $p=.83$).

Design. A 2 (age group: younger vs. older) x 2 (retrieval condition: control vs. SC) x 4 (detail type: episodic vs. incorrect vs. source error vs. semantic) mixed design was employed with repeated measures on the second and third factors. In the SC condition participants viewed all of the pictures from their own device in forward temporal order, while in the control condition participants did not see any pictures. Memory performance was measured as the number of details produced in a verbal recall attempt. More information about the types of details is presented in the *Coding strategy* section, below.

Materials and procedure

---

in a third sample of participants who took part in an unrelated experiment (not yet published). We examined the relationship between measures in that sample. Pearson's correlations showed a strong positive relationship between NART and WASI scores ($r=.82$, $df=38$, $p<.0005$). We also calculated the mean difference between WASI and NART IQ estimates and found that the NART tended to slightly underestimate IQ compared to the WASI. For younger adults, WASI scores were *M*=2.22 (*SD*=5.95) points higher than NART scores, and for older adults WASI scores were *M*=2.28 (*SD*=8.90) points higher. A t-test showed that the difference in scores was equivalent for both groups ($t(38)=.03$, $p=.98$).

Encoding session. The encoding session for Experiment 2 was similar to that of Experiment 1, but took place over two rooms instead of three, to reflect the reduced number of review conditions. Participants were placed in mixed-age groups of up to 10 individuals (self-selected by their availability on the choice of test dates), and each participant was provided with a SC, which was worn for the duration of the encoding session. Each of the two rooms contained three novel tasks, which were different from those used in Experiment 1, and were designed to provide more nonvisual material to be recalled at test. Participants remained in the same groups throughout the event and visited each room in turn, with the same experimenter acting as task leader in both rooms. All participants visited the two rooms in the same order. In Room 1, the tasks involved matching criminal mugshots to their crimes, folding an origami pigeon, and completing a problem solving task in which nine dots must be joined by drawing four connected lines (Maier, 1930). In Room 2, the tasks involved tasting chocolate and identifying the flavour, completing a word-search, and telling two personal truths and one lie.

In order to control for any differences in room memorability, counterbalancing was implemented at the review stage such that half of the participants saw photographs for Room 1 and the other half saw photographs for Room 2. At the end of the encoding session, participants returned their cameras and the stimuli were prepared for the retrieval session.

Retrieval session. The retrieval session took the form of a one-to-one interview 13-15 days after the encoding session. Participants were first asked to recall the room for which they did not see photographs. The retrieval instructions asked participants to recall as much as possible about the room, and specifically to describe the tasks, social interactions and thoughts or feelings they had at the time, while being as specific as possible and reporting details even if they seemed insignificant. To avoid excluding a large amount of data from the analysis, participants were not explicitly asked to recall visual information. Following recall of the non-reviewed room, participants next reviewed the SC pictures for the remaining room. Images were presented serially on a laptop computer, and were reviewed in forward temporal order. Participants self-paced the review by pressing the forward arrow on the keyboard to move through the sequence. No further instructions were given for review, and any memories the participants generated during review were not recorded. Immediately after reviewing the second room, participants were asked to recall the room under the same instructions as above.

Coding strategy. The recall sessions were audio recorded, transcribed and coded for episodic and semantic details, incorrect details, and source errors. As in Experiment 1, a master coding sheet was created for the scripted aspects of the event, to ensure that coding was consistent across all participants. Due to the idiosyncrasy of event recall it was not possible to anticipate memory for the non-scripted information, however care was taken to code recall for both types of information at the same grain of detail. Episodic details were those specific to the event, and included actions (e.g. *folding an origami pigeon*), visual details (e.g. *using blue origami paper*), interactions with other individuals, auditory information (e.g. *hearing road works outside the window*), conceptual details (e.g. crimes committed), and thoughts and feelings that the participant had at the time (see *Appendix C* for a coded example). Incorrect details were of a similar nature, but described things that either did not happen, or that were reported inaccurately (e.g. remembering five types of chocolate instead of four). Details that were conceptually correct but were reported as part of the wrong room were counted as source memory errors, while memory details that were not bound to a specific time and place (e.g. “I like chocolate”), were coded as semantic. As in Experiment 1, the reliability of the coding strategy was measured by comparing the original coding with that of three additional raters, who each rated the same subset of transcripts

(n=6). A two-way random intraclass correlation showed good agreement (r_{ic}=.99; 95% CI = .95, 1.00).

Results

The number of each type of detail (episodic, incorrect, source error, semantic) recalled in each condition is presented in *Table 3*. Source errors and semantic details were not included in subsequent analyses due to the low numbers recalled by both groups, which were almost at floor. Similarly to Experiment 1, we analysed episodic and incorrect details separately.

We began by analysing the effect of age and SC review on episodic recall in a 2x2 ANOVA. There was a significant main effect of age (F(1,22)=9.09, p=.006, η_p^2=.29, MSE=320.67), in which younger adults recalled more episodic details (M=35.75, SD=17.91) than older adults (M=20.167, SD=17.91). There was also a main effect of condition (F(1,22)=15.61, p=.001, η_p^2=.42, MSE=56.64), which showed that more episodic details were recalled after SC-review (M=32.25, SD=14.85) than in the control condition (M=23.67, SD=12.54). There was no interaction between age and condition (F(1,22)=.12, p=.73, η_p^2=.01, MSE=56.64). [9] A similar analysis was carried out for incorrect details, but found no effect of age (F(1,22)=2.57, p=.12, η_p^2=.11, MSE=5.48) or condition (F(1,22)=.10, p=.75, η_p^2=.01, MSE=3.25), and no age by condition interaction (F(1,22)=.03, p=.87, η_p^2<.01, MSE=3.25).[10]

[9] In the original, non-matched sample the pattern of results for episodic recall was similar, apart from the absence of a main effect of age (F(1,34)=.36, p=.55, η_p^2=.01, MSE=640.29). The main effect of condition was significant (F(1,34)=26.70, p<.0005, η_p^2=.44, MSE=68.44), but there was no age by condition interaction (F(1,34)=.34, p=.56, η_p^2=.01, MSE=68.44).

[10] The number of incorrect details recalled by the full sample was not affected by age (F(1,34)=1.25, p=.27, η_p^2=.04, MSE=7.02) or condition (F(1,34)=1.35, p=.25, η_p^2=.04, MSE=644.30) and there was no interaction (F(1,34)=.31, p=.58, η_p^2=.01, MSE=4.30).

We next looked at the number of recall errors as a proportion of total recall output; these data are presented in *Table 3*. A 2x2 ANOVA showed that the proportion of incorrect details was not affected by SC condition ($F(1,22)=.29$, $p=.59$, $\eta_p^2=.01$, $MSE=.004$) or age ($F(1,22)=.61$, $p=.44$, $\eta_p^2=.03$, $MSE=.003$), and there was no interaction ($F(1,22)=.31$, $p=.59$, $\eta_p^2=.01$, $MSE=.004$). Together these results suggest that SC supports recall of episodic event details without increasing the recall of inaccurate event information.[11]

[Table 3 about here]

Nonvisual details.

As in Experiment 1, it is possible that some of the additional episodic details participants reported were seen in the SC image sequence a few minutes previously, rather than details that the participant recalled from the encoding session two weeks earlier. To investigate this possibility, we excluded visual episodic details. Again, the decision to remove all visual details was a conservative measure because it was likely that some of the visual information was not present in each participant's image sequence. A 2x2 ANOVA analysed the effect of age and condition on the recall of episodic nonvisual details only. There was a main effect of condition ($F(1,22)=25.46$, $p<.0005$, $\eta_p^2=.54$, $MSE=37.48$), in which the SC condition was associated with more nonvisual episodic recall ($M=27.63$, $SD=15.12$) than the control condition ($M=18.71$, $SD=13.20$). There was also a main effect of age ($F(1,22)=7.10$, $p=.01$, $\eta_p^2=.24$, $MSE=289.39$), in which younger adults ($M=29.71$, $SD=14.86$) recalled more nonvisual details than older adults ($M=16.63$, $SD=8.27$); age and condition did not interact ($F(1,22)=.57$, $p=.46$, $\eta_p^2=.03$, $MSE=37.48$).[12]

[11] The proportion of incorrect details was not affected by condition ($F(1,33)=.03$, $p=.86$, $\eta_p^2<.01$, $MSE=.004$) or age group ($F(1,33)=.59$, $p=.45$, $\eta_p^2=.02$, $MSE=.006$), and there was no interaction ($F(1,33)=.07$, $p=.80$, $\eta_p^2<.01$, $MSE=.004$).

[12] In the full sample, more nonvisual details were recalled after SC review compared to in the control condition ($F(1,34)=35.62$, $p<.0005$, $\eta_p^2=.51$, $MSE=49.27$), but there was no effect of age on nonvisual recall ($F(1,34)=.72$, $p=.40$, $\eta_p^2=.02$, $MSE=416.36$), and no age by condition interaction ($F(1,34)=.84$, $p=.37$, $\eta_p^2=.02$, $MSE=49.27$).

Examination of the mean number of episodic non-visual details recalled by older adults after viewing SC images (M=20.42, SD=11.06) showed that they recalled a similar number of details as younger adults recalled without SC support (M=24.58, SD=15.17), which suggests that SC can compensate for age-related episodic memory deficits by prompting recall of material that was not available in the images.

Repeat participants.

As noted above, six older adults and two younger adults who participated in Experiment 2 also participated in Experiment 1. Although the experimental procedure was fully explained to all participants at the start of the study, and the design of the study changed substantially in Experiment 2, it is possible that the prior experience of these participants affected the results. However, comparison of the mean number of episodic details recalled by participants in Experiment 2 who did (M_{young}=19.50, SDy_{oung}=16.26; M_{older}=23.83, SD_{older}=9.58) and did not (M_{young}=26.40, SD_{young}=15.46; M_{older}=22.85, SD_{older}=22.93) take part in Experiment 1 suggested that any effect of prior experience was likely to be minimal. Moreover, in the matched sample half of the repeat participants were excluded, leaving only two older adults and one young adult who participated in both experiments.

Discussion

The results of Experiment 2 showed that both older and younger adults recalled more specific event details after reviewing SC images of the event, compared to a baseline condition in which recall was not supported by prior review. These results replicated the findings in Experiment 1, and extended the findings from our earlier work on everyday events (Mair et al., 2017) to demonstrate a SC benefit for recall of controlled staged events. The SC effect was observed even when the analysis excluded visual details reported by participants, which suggests that SC cues recollection of details from the original event, and not just recognition of information seen in the SC photographs (Silva et al., 2016). Moreover, older adults' recall after reviewing SC images was as detailed as younger adults' unsupported recall, which reveals potential for the use of SC-like devices in everyday life as relatively cheap and accessible memory aids for older adults, for the retrieval of specific events. However, it should be noted that the older adults who participated in both Experiment 1 and Experiment 2 were predominantly a "young-old" group (i.e. aged 65-75), and it is possible that and "old-old" group would respond differently to a technological memory aid.

General discussion

The studies presented here demonstrate the ability of SC to support recollection of recently experienced events in older and younger adults, without increasing the amount of inaccurate recall. These results were based on the use of SC images as retrieval cues, whereby images were uploaded to a computer, and participants were presented with the full set of available pictures. It should be pointed out that this is only one of a number of ways SC can be used, and as yet it is unclear how to provide the most effective memory support. Indeed, there appear to be some instances in which SC does not provide much recall advantage. For example, a recent study investigated 144 young adults' recall of unusual actions experienced at particular locations during a staged walk around a university campus. Event review took place minutes after the event itself, and recall was tested after an interval of one week. In that study there was no benefit of reviewing SC images of the walk compared to either reviewing written prompts, or an unsupported control condition in which the "review" was a free recall attempt (Seamon et al., 2014). This suggests that the use of SC as a consolidation support in healthy individuals may be of limited value.

In Experiment 1 we found little effect of randomising the order of SC images within a cue sequence, suggesting that when all images are presented the sheer amount of information available in the cue may be the primary determinant of SC's success as a retrieval aid

(Hodges et al., 2011). It is possible that, given the small sample size, a small effect of temporal order was not detected due to insufficient power, although any such effect would be of limited applied value. However, if devices such as SC are to be used on a regular basis, or for longer durations, it is of practical importance to develop ways to reduce the amount of time spent reviewing the images while maintaining the mnemonic benefit, and for this reason it is important to understand the mechanism of the SC retrieval benefit It may be that the compatibility of SC image sequences with normal memory (Hodges et al., 2011) is important when those image sequences contain less information. Future research should aim to address these complex issues.

References

Anderson, S. J. & Conway, M. A. (1993). Investigating the structure of autobiographical memories. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19 (5)*, 1178—1196.

Badham, S. P. & Maylor, E. A. (2015). What you know can influence what you are going to know (especially for older adults). *Psychonomic Bulletin and Review, 22,* 141—146.

Badham, S. P., Hay, M., Foxon, N., Kaur, K., & Maylor, E. A. (2016). When does prior knowledge disproportionately benefit older adults' memory? *Aging, Neuropsychology, and Cognition, 23 (3),* 338—365.

Berry, E., Kapur, N., Williams, L., Hodges, S., Watson, P., Smyth, G., Srinivasan, J., Smith, R., Wilson, B. & Wood, K. (2007). The use of a wearable camera, SenseCam, as a pictorial diary to improve memory in a patient with limbic encephalitis: A preliminary report. *Neuropsychological Rehabilitation, 17 (4/5),* 582-601.

Birch, L. S. & Davidson, K. M. (2007). Specificity of autobiographical memory in depressed older adults and its relationship with working memory and IQ. *British Journal of Clinical Psychology, 46,* 175-186.

Bright, P., Jaldow, E. & Kopelman, M. D. (2002). The National Adult Reading Test as a measure of premorbid intelligence: A comparison with estimates derived from demographic variables. *Journal of the International Neuropsychological Society, 8,* 847-854.

Browne, G., Berry, E., Kapur, N., Hodges, S., Smyth, G., Watson, P., & Wood, K. (2011). SenseCam improves memory for recent events and quality of life in a patient with memory retrieval difficulties. *Memory, 19 (7),* 713-722.

Cochrane, W. G. (1957). Analysis of covariance: Its nature and uses. *Biometrics, 13 (3),* 261-281.

Conway, M. A. & Pleydell-Pearce, C. W. (2000). Construction of autobiographical memories in the Self-Memory System. *Psychological Review, 107 (2),* 261-288.

Finley, J. R., Brewer, W. F. & Benjamin, A. S. (2011). The effects of end-of-day picture review and a sensor-based picture capture procedure on autobiographical memory using SenseCam. *Memory, 19 (7),* 796-807.

Folstein, M. F., Folstein, S. E. & McHugh, P. R. (1975). "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12 (3),* 189-198.

Galton, F. (1879). Psychometric experiments. *Brain, 2 (2),* 149-162.

Hodges, S., Berry, E. & Wood, K. (2011). SenseCam: A wearable camera that stimulates and rehabilitates autobiographical memory. *Memory, 19 (7),* 685-696.

Kalnikaité, V., Sellen, A., Whittaker, S., & Kirk, D. (2010). *Now let me see where I was: Understanding how lifelogs mediate memory.* Paper presented at the Proceedings of 28th International Conference on Human Factors in Computing Systems - CHI '10.

Koutstaal, W., Schacter, D. L., Johnson, M. K., & Galluccio, L. (1999). Facilitation and impairment of event memory produced by photograph review. *Memory & Cognition, 27 (3),* 478-493.

Koutstaal, W., Schacter, D. L., Johnson, M. K., Angell, K. E., & Gross, M. S. (1998). Post-event review in older and younger adults: Improving memory accessibility of complex everyday events. *Psychology and Aging, 13 (2),* 277-296.

Kvavilashvili, L., Mirani, J., Schlagman, S., Erskine, J. A. K. & Kornbrot, D. E. (2010). Effects of age on phenomenology and consistency of flashbulb memories of September 11 and a staged control event. *Psychology and Aging, 25(2),* 391-404.

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

Light, L. L., & Anderson, P. A. (1983). Memory for scripts in young and older adults. *Memory & Cognition, 11 (5),* 435—444.

Loveday, C. & Conway, M. A. (2011). Using SenseCam with an amnesic patient: Accessing inaccessible everyday memories. *Memory, 19 (7),* 697-704.

Maier, N. R. F. (1930). Reasoning in humans. I. On direction. *Journal of Comparative Psychology, 10 (2),* 115-143.

Mair, A., Poirier, M., & Conway, M. A. (2017). Supporting older and younger adults' memory for recent everyday events: a prospective sampling study using SenseCam. *Consciousness and Cognition*, DOI: 10.1016/j.concog.2017.02.008

Miller, G. A. & Chapman, J. P. (2001). Misunderstanding analysis of covariance. *Journal of Abnormal Psychology, 110 (1),* 40-48.

Mueller-Johnson, K. & Ceci, S. J. (2004). Memory suggestibility in older adults: Live event participation and repeated interview. *Applied Cognitive Psychology, 18,* 1109—1127.

Nelson, H. E. (1982). *National Adult Reading Test (NART): For the assessment of premorbid intelligence in patients with dementia.* NFER-Nelson, Windsor, UK.

Pauly-Takacs, K., Moulin, C. J. A., & Estlin, E. J. (2011). SenseCam as a rehabilitation tool in a child with anterograde amnesia. *Memory, 19 (7),* 705-712.

Piolino, P., Coste, C., Martinelli, P., Macé, A.-L., Quinette, P., Guillery-Girard, B. & Belleville, S. (2010). Reduced specificity of autobiographical memory and aging: Do the executive and feature binding functions of working memory have a role? *Neuropsychologia, 48,* 429-440.

Piolino, P., Desgranges, B., Benali, K. & Eustache, F. (2002). Episodic and semantic remote autobiographical memory in ageing. *Memory, 10 (4),* 239-257.

Piolino, P., Desgranges, B., Clarys, D., Guillery-Girard, B., Taconnat, L., Isingrini, M. & Eustache, F. (2006). Autobiographical memory, autonoetic consciousness and self-perspective in aging. *Psychology and Aging, 21 (3),* 510-525.

Rosen, V. M., Caplan, L., Sheesley, L., Rodriguez, R., & Grafman, J. (2003). An examination of daily activities and their scripts across the adult lifespan. *Behavior Research Methods, Instruments, & Computers, 35 (1),* 32—48.

Seamon, J. G., Moskowitz, T. N., Swan, A. E., Zhong, B., Golembeski, A., Liong, C., Narzikul, A. C., & Sosan, O. A. (2014). SenseCam reminiscence and action recall in memory-unimpaired people. *Memory, 22 (7),* 861-866.

Sellen, A., Fogg, A., Aitken, M., Hodges, S., Rother, C. & Wood, K. (2007). *Do life-logging technologies support memory for the past? An experimental study using SenseCam.* Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, San Jose, California, USA, pp. 81-90. DOI: 10.1145/1240624.1240636

Silva, A. R., Pinho, M. S., Macedo, L., & Moulin, C. J. A. (2016). A critical review of the effects of wearable cameras on memory. *Neuropsychological Rehabilitation*, DOI: 10.1080/09602011.2015.1128450.

Silva, A. R., Pinho, S., Macedo, L. M. & Moulin, C. J. (2013). Benefits of SenseCam review on neuropsychological test performance. *American Journal of Preventative Medicine, 44 (3),* 302-307.

St. Jacques, P. L. & Schacter, D. L. (2013). Modifying memory: Selectively enhancing and updating personal memories for a museum tour by reactivating them. *Psychological Science, 24 (4),* 537-543.

St. Jacques, P. L., Montgomery, D. & Schacter, D. L. (2015). Modifying memory for a museum tour in older adults: Reactivation-related updating that enhances and distorts memory is reduced in ageing. *Memory, 23 (6),* 876-887.

Umanath, S. & Marsh, E. J. (2014). Understanding how prior knowledge influences memory in older adults. *Perspectives on Psychological Science, 9 (4),* 408-426.

Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence*. The Psychological Corporation: Harcourt Brace & Company. New York, NY.

West, R. L. & Stone, K. R. (2014). Age differences in eyewitness memory for a realistic event. Journals of Gerontology, Series B: *Psychological Sciences and Social Sciences, 69 (3),* 338-347.

Woodberry, E., Browne, G., Hodges, S., Watson, P., Kapur, N., & Woodberry, K. (2015). The use of a wearable camera improves autobiographical memory in patients with Alzheimer's disease. *Memory, 23 (3),* 340-349.

Yesavage, J. A., Brink, T. L., Rose, T. L., Lum, O., Huang, V., Adey, M. & Leirer, V. O. (1983). Development and validation of a geriatric depression screening scale: A preliminary report. *Journal of Psychiatric Research, 17,* 37-49.

Table 1

Mean number of details recalled by young and older adults at T1

	Young adults		Older adults	
Detail type	*M*	*SD*	*M*	*SD*
Total episodic	21.90	9.68	13.19	7.48
Total incorrect	2.21	1.65	1.91	1.13
Total source error	1.02	2.33	2.23	2.08
Proportion incorrect	.08	.04	.13	.08
Proportion source error	.05	.14	.14	.12

Table 2

Mean number of new details recalled at T2

		Young adults		Older adults	
Detail type	Condition	*M*	*SD*	*M*	*SD*
Episodic	Control	3.80	4.33	3.05	3.24
	Random	6.79	2.91	8.53	6.72
	Forward	8.47	4.36	6.68	4.92
Incorrect	Control	.87	1.30	1.00	1.45
	Random	1.07	1.07	1.47	1.43
	Forward	1.33	1.29	1.26	1.15
Proportion incorrect	Control	.12	.23	.24	.31
	Random	.14	.13	.20	.19
	Forward	.12	.11	.21	.19
Nonvisual	Control	2.33	3.60	1.37	1.74
	Random	3.07	2.27	2.74	2.40
	Forward	4.27	3.43	2.53	2.63

Table 3

Mean number of details recalled by young and older adults at baseline and after reviewing SenseCam images

	Young adults				Older adults			
	Baseline		SenseCam		Baseline		SenseCam	
Detail type	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Episodic	31.08	14.89	40.42	15.95	16.25	9.59	24.08	13.66
Incorrect	2.42	3.50	2.33	1.37	1.42	1.24	1.17	1.34
Source error	.17	.58	.17	.58	.83	2.59	.00	.00
Semantic	.58	.90	.92	1.31	.92	1.38	1.83	2.52
Prop. incorrect	.06	.06	.06	.06	.08	.06	.06	.07
Nonvisual	24.58	15.17	34.83	15.57	12.83	7.66	20.42	11.06
Prompted	15.33	6.46	13.83	11.53	12.92	5.99	9.58	8.22

Note. Prop. incorrect = incorrect details expressed as a proportion of the total recalled details; Prompted = details added to memory narratives after specific verbal prompts.

Appendix A

Questionnaire items used in Experiment 1

Question number	Question
1	You had two minutes to look around the [red/blue/green] room. Please describe everything you remember seeing.
2	Please describe the task leader.
3	Do you remember any of the general knowledge questions from the [red/blue/green] room? Please describe anything that you remember about this task, including any other information you received from the task leader.
4	Do you remember any of the flags from the [red/blue/green] room? Please describe anything that you remember about this task, including any other information you received from the task leader.
5	Do you remember what the problem-solving task involved in the [red/blue/green] room? Please describe what your group had to do and the roles each person took to complete the task.

Appendix B

Encoding session tasks employed in Experiment 1

Room	General knowledge	Name the flag	Problem solving
Red	1. Which animal were the Canary Islands named after? 2. How many miles of blood vessels are there in the human body? 3. What is strange about the jellyfish species *Turritopsis Dohrnii*?	1. Morocco 2. Denmark 3. Turkey	*Lego task* One participant was a describer, the other one or two participants were builders. The describer sat behind a screen with a lego model, which he/she described to the builders. The builders had to make a replica of the model from the describer’s instructions, without seeing the original.
Blue	1. What is the circumference of the Earth, at the equator, in miles? 2. In 1923, jockey Frank Kayes won a race at Belmont Park in New York. What was strange about his victory? 3. What are Australian Mist, Cornish Rex, Scottish Fold and Turkish Van types of?	1. Kazakhstan 2. Somalia 3. Greece	*Match stick puzzle* Participants presented with an array of sixteen match sticks arranged in two squares. They worked together to move four match sticks to create three squares.
Green	1. How many towns in Great Britain are called Newport or have Newport in the name? 2. What are the small circles of paper called that are cut out by a hole-puncher? 3. What is the world’s biggest island?	1. Pakistan 2. Jamaica 3. Nigeria	*Word wheel* Participants provided with a nine-letter word wheel and asked to make as many words as possible with their group. Words had to be at least four letters long, and there was one nine-letter word to find.

Appendix C

Example response from Experiments 1 and 2, coded for visual/nonvisual episodic detail

Experiment 1

The problem solving task involved a wheel with different letters in it **[V]** We had to come up with as many words as possible **[NV]** which had a minimum of 4 letters **[NV]** There was a 9 letter word which used all of the letters which we had to try and find **[NV]** It was quite easy to come up with words **[NV]** and we both worked well together **[NV]** The 9 letter word **[R]** which we did not guess **[NV]** was 'Celebrity' **[NV]** I remember thinking that I thought the word began with C **[NV]**

Experiment 2

Okay so we had to taste chocolate **[V]** and I think there was a crystallised ginger chocolate **[NV]** … there was a strawberry one **[NV]** and I put raspberry **[NV]** and there was a chilli one **[NV]** which I got right **[NV]** and then there was a… I'll come back to that one **[X]** But anyway, I got like two and a half right **[NV]** And then we sat down **[V]** and we had to tell lies about ourselves **[NV]** So I said that… yeah nobody believed **[NV]** that I'd got a physics degree **[NV]** and I said that I was Romanian **[NV]** and the other one was my age **[NV]** Someone else said "I can't use computers" **[NV]** and then it turned out that she worked with computers **[NV]** and she found it really funny **[NV]**

Note. V= correct episodic visual detail (i.e. may be visible within SenseCam image sequence); NV = episodic nonvisual detail (i.e. could not be visible within a SenseCam image sequence); R = repeated detail; X = utterance not counted as a memory detail