

1 **Facilitating recall and particularization of repeated events in adults**
2 **using a multi-method interviewing format**

3 Feni Kontogianni^{a*}, Eva Rubinova^a, Lorraine Hope^a, Paul J. Taylor^b,
4 Aldert Vrij^a, and Fiona Gabbert^c

5 ^a *Department of Psychology, University of Portsmouth, Portsmouth, United Kingdom;*

6 ^b *Department of Psychology, Lancaster University, Lancaster, United Kingdom &*
7 *University of Twente, Enschede, The Netherlands;*

8 ^c *Department of Psychology, Goldsmiths University of London, United Kingdom*

9 *Corresponding author¹: Feni Kontogianni, Department of Psychology, University of
10 Portsmouth, Portsmouth, PO1 2DY (email: feni.kontogianni@port.ac.uk)

11

12 Short title: ADULTS' RECALL OF REPEATED EVENTS

13

14 Word count: 8092

15

16 *This is an Accepted Manuscript of an article published by Taylor & Francis in Memory*
17 *on 10.3.21, available online: <http://dx.doi.org/10.1080/09658211.2021.1903508>.*

18

19 Supplemental data for this article can be accessed through the following link:

20 <https://doi.org/10.1080/09658211.2021.1903508>

21

22 Data and analysis code used in this research are available on the Open Science

23 Framework and can be accessed at: <https://osf.io/4mcsa/>

24

¹ The corresponding author's affiliation has changed: Department of Psychology, University of
Winchester, Winchester, SO22 4NR (email: feni.kontogianni@winchester.ac.uk)

1 **Facilitating recall and particularization of repeated events in adults**
2 **using a multi-method interviewing format**

3 Reports about repeated experiences tend to include more schematic information
4 than information about specific instances. However, investigators in both forensic
5 and intelligence settings typically seek specific over general information. We
6 tested a multi-method interviewing format (MMIF) to facilitate recall and
7 particularization of repeated events through the use of the self-generated cues
8 mnemonic, the timeline technique, and follow-up questions. Over separate
9 sessions, 150 adult participants watched four scripted films depicting a series of
10 meetings in which a terrorist group planned attacks and planted explosive
11 devices. For half of our sample, the third witnessed event included two deviations
12 (one new detail and one changed detail). A week later, participants provided their
13 account using the MMIF, the timeline technique with self-generated cues, or a
14 free recall format followed by open-ended questions. As expected, more
15 information was reported overall in the MMIF condition compared to the other
16 format conditions, for two types of details, correct details, and correct gist details.
17 The reporting of internal intrusions was comparable across format conditions.
18 Contrary to hypotheses, the presence of deviations did not benefit recall or source
19 monitoring. Our findings have implications for information elicitation in applied
20 settings and for future research on adults' retrieval of repeated events.

21 Keywords: repeated events; memory retrieval; timeline technique; self-generated
22 cues; deviation

23

24

25

26

27

28

1 Witnesses, victims, or sources may be interviewed about a series of events that have
2 occurred repeatedly over a period of time, such as, domestic violence, sexual abuse,
3 industrial accident investigations, or meetings of a criminal gang. In investigations
4 involving repeated events, interviewees will likely need to retrieve and report details of
5 a specific incident (i.e., particularization; see p. 203, Brubacher & Roy, 2014; p. 67,
6 Powell et al., 2007). To date, numerous studies have examined how this can be achieved
7 when interviewing children about repeated events of child abuse (Brubacher et al.,
8 2014; Woiwod et al., 2019). However, there is little forensic research on adults’
9 memory of repeated events (e.g., MacLean et al., 2018) and, more particularly, on
10 techniques that might be used to effectively elicit information for specific instances
11 (e.g., Leins et al., 2014; Theunissen et al., 2017; Willen et al., 2015).

12 Memory for repeated experiences differs from memory for unique experiences
13 in various ways (Price & Connolly, 2013). Across a series of repeated events, some
14 details are recurring and thus form a routine (e.g., every meeting starts with the leader
15 of the terrorist group describing a plan). Other details vary across instances (e.g., a
16 different target is selected for each attack). Because of repeated exposure, memory for
17 recurring details is stronger than memory for variable details. Similarly, memory for the
18 overall routine across events is stronger than for a specific incident within the series
19 (McNichol et al., 1999). Consequently, interviewees are likely to under-report
20 information about specific instances.

21 Interviewees are often required to describe instances with precision, such as by
22 reporting dates or times, or by identifying perpetrators of specific actions. However,
23 discriminating between repeated events is a challenging task, as there is a high
24 likelihood of memory interference, particularly when the events are similar (Farrar &
25 Boyer-Pennington, 1999; Lindsay & Johnson, 1989). Interviewees might attribute a

1 detail to the wrong incident (i.e., source misattribution) or they might remember what
2 occurred but not when it occurred (Johnson et al., 1993); both of which can negatively
3 affect the interviewee's credibility and the efficacy of the investigation (Brubacher et
4 al., 2014; Weinsheimer et al., 2017).

5 There are distinct challenges in remembering and reporting repeated events that
6 are not fully addressed by current information gathering protocols. Drawing from a rich
7 theoretical framework on the representation of repeated events in memory, the current
8 study tests the effectiveness of a multi-method interviewing format (MMIF), which
9 includes: i) the self-generated cues mnemonic (SGC; Kontogianni et al., 2018; Wheeler
10 & Gabbert, 2017); ii) the timeline technique (Hope et al., 2013; Hope et al., 2019); and,
11 iii) the use of follow-up open-ended questions (Kontogianni et al., 2020). These
12 elements have been previously tested separately or in combination in research
13 examining memory for single events. To examine their combined effectiveness, we test
14 them against two comparison groups: one intermediate group where participants used
15 the self-generated cues and the timeline technique (SGC-Timeline), and a control group
16 where participants used a free recall format followed by open-ended questions.

17 **Repeated events: Schema and Fuzzy-Trace Theory**

18 Repeated events are thought to be represented in memory as parts of an overarching
19 higher-order knowledge structure that is characterized by a common theme, referred to
20 as a schema (Ahn et al., 1992; Brewer & Nakamura, 1984). Schema theory suggests that
21 one experience suffices to begin building a script about what typically occurs in an
22 event. The experience of additional similar occurrences informs a more elaborate
23 schematic representation and shapes our expectations about future occurrences (Ahn et
24 al., 1992; Farrar & Boyer-Pennington, 1999; Hudson et al., 1992). In line with the

1 spreading activation theory of memory (Anderson, 1983), repeated exposure to
2 recurring details forms strong associated traces within the memory network which in
3 turn increases the probability of their retrieval (e.g., Hudson et al., 1992). Over the
4 course of a repeated event, some *variations* (i.e., predictable alternatives of certain
5 recurring actions) are likely to occur (Abelson, 1981). Compared to recurring details
6 that are stable and characterise the general routine of the events, variations are thought
7 to be absorbed by the script over time and, thus, become less likely to be retrieved
8 (Abelson, 1981; Schank & Abelson, 1977). This notion aligns with the idea that, over
9 time, the content of repeated instances can become part of semantic memory in an
10 abstracted form (e.g., Brewer & Nakamura, 1984).

11 However, there is evidence that *deviations* from the script, which are atypical
12 and unpredictable details, can be particularly memorable (Abelson, 1981). To the extent
13 that deviations are schema-inconsistent, research suggests that they are likely to be
14 recalled because they attract attention and require increased resources to be integrated to
15 the script, therefore comprising a strong trace in the memory network (Anderson, 1983;
16 Brewer & Nakamura, 1984). In other words, deviations are more memorable because
17 they violate the script and are thus distinctive and salient (Davidson, 2006; see also
18 Cohen & Java, 1995; Means & Loftus, 1991). There is also evidence that deviations can
19 improve source monitoring because they serve as tags for specific instances (*“script*
20 *pointer plus tag” hypothesis*; Graesser et al., 1980). Recent studies that have
21 manipulated the presence of deviations in a target instance of repeated events show that
22 deviations can improve recall for that instance (targeted effect; MacLean et al., 2018),
23 or even for all the events (general effect; Connolly et al., 2016; MacLean et al., 2018). If
24 deviations from the script improve recall and particularization of repeated events, their
25 occurrence can have implications for information elicitation.

1 Fuzzy-Trace Theory (FTT; Brainerd & Reyna, 1990) offers a similar
2 conceptualization to schema theory regarding the retrieval of repeated events. Fuzzy-
3 Trace Theory suggests that experiences are encoded and stored in memory in the form
4 of two traces: gist, which in the context of repeated events refers to an overall
5 understanding of what typically occurs (e.g., general recurring details); and verbatim,
6 which represent instance-specific details (e.g., sources of variations; deviations).
7 Research shows that gist and verbatim traces are retrieved separately via different cues,
8 and that verbatim traces are more sensitive to forgetting than gist traces, so that memory
9 over time tends to rely on gist (*principle of retrieval dissociation*; Brainerd & Reyna,
10 2001; 2004). In sum, although they differ regarding the recall of deviations, both FTT
11 and schema theory suggest that specific details of separate instances are less likely to be
12 accessed over time compared to the general routine of repeated events. This is true
13 either because specific details fade out from memory more rapidly, or because they
14 become absorbed by the memory of the routine itself.

15 **Interviewing techniques to facilitate recall and reporting**

16 In line with the theoretical work reviewed above, research has tested various cues to
17 facilitate recall about both the general routine and specific details of repeated events –
18 although cues are usually tested discretely rather than in combination. Most methods
19 capitalize on the thematical and temporal organisation of autobiographical memories,
20 according to which specific events are thought to be hierarchically nested within
21 summarised, and extended events (Conway & Pleydell-Pearce, 2000). For instance,
22 research suggests that inquiring about the overall number of witnessed events first,
23 before asking about specific instances leads to more elaborate reporting (Connolly &
24 Gordon, 2014). Another recommendation is that open and directive prompts are used, in
25 that order, to inquire about the frequency of specific actions and variations (Brubacher

1 et al., 2014). In survey methodology, visual timelines have been effectively used to
2 elicit information about major thematic events (e.g., relationships, work) over prolonged
3 time periods (*event history calendars*; Belli, 1998; Van der Vaart & Glasner, 2007;
4 Yoshihama et al., 2005). With respect to more mundane events, Means and Loftus
5 (1991) interviewed participants about repeated health care visits, and found that asking
6 participants to think about specific elements of each visit (e.g., type of doctor, weather
7 etc.) and then construct a personal timeline improved the amount of recall and dating
8 accuracy of separate instances relative to reporting following a free recall. More
9 recently, Leins et al. (2014) found that the use of a timeline can benefit the reporting of
10 details of family gatherings when administered with various mnemonics, such as the
11 ‘family tree mnemonic’ (cf. unaided recall), as participants may identify meetings in
12 relation to other temporal markers. For instance, participants were prompted with
13 derived cues that asked about normative events likely shared by members of the same
14 social background (e.g., Thanksgiving). They were then asked to report why such
15 gatherings occurred in their family, and these personally relevant reasons were used as
16 cues to prompt retrieval. A similar cuing technique was used in Willen et al. (2015),
17 where context-specific cues were derived from the most salient details that participants
18 remembered from a series of dental visits. When other participants used the context-
19 specific cues from a similar perspective to theirs, they reported more specific instance
20 details than when they used general cues, such as times and dates.

21 With respect to facilitating recall and particularization of details, the current
22 research investigated how the use of self-generated cues (Kontogianni et al., 2018;
23 Wheeler & Gabbert, 2017), the timeline technique (Hope et al., 2013), and open-ended
24 questions (Kontogianni et al., 2020) can be used in conjunction as a multi-method
25 interviewing format. The use of the timeline should effectively cue the retrieval of

1 details that are temporally and thematically associated in the context of repeated events
2 (Belli, 1998). To date, the timeline technique has been found to enhance the reporting of
3 accurate details, sequential information, and information about attributions of actions to
4 multiple perpetrators in single events; both as a stand-alone technique (Hope et al.,
5 2013) and in conjunction with the self-generated cues mnemonic (Kontogianni et al.,
6 2018). To promote ‘top-down’ retrieval, we used a modified timeline so that
7 interviewees are first asked to outline the number of the witnessed events before being
8 asked to use the self-generated cues and describe each instance on a separate format
9 (see also Hope et al., 2019).

10 Self-generated cues prompt interviewees to list the most salient details from
11 each instance, and thus facilitate the retrieval of closely associated memories (Wheeler
12 & Gabbert, 2017). Previous research has shown that when participants are asked to
13 generate their own cues during encoding of target items, cues with distinctive properties
14 are more specific than experimenter-generated general descriptions (Mantyla & Nilsson,
15 1988; Tullis & Benjamin, 2015). Anderson and Conway (1993) have also shown that
16 participants tend to list distinctive details about previously experienced events first,
17 followed by other thematically related details. Thus, evidence shows that reliable cues
18 both reinstate the context of the experienced event (Tulving & Thomson, 1973) and
19 provide diagnostic information about specific memories (*principle of cue overload*; Goh
20 & Lu, 2012; Nairne, 2002). Based on previous applied research, we expected that the
21 use of self-generated cues would facilitate the discrimination of specific instances
22 thereby improving recall while reducing source confusion (Brubacher et al., 2018;
23 Willen et al., 2015).

24 Interference between instances might lead to source intrusions and the reporting
25 of inconsistent details (e.g., confusing a perpetrator’s actions in one instance based on

1 what was witnessed in another instance) and omission errors (see also Lindsay, 2014).
2 Indeed, practitioners often follow-up on an interviewee's account to clarify what has
3 been reported and to address information gaps (Shepherd & Griffiths, 2013). There is
4 evidence that the use of open questions that prompt separate instances in depth, rather
5 than in breadth, is associated with the reporting of more specific (cf. general routine)
6 details (Brubacher et al., 2012). In addition, their use could prompt source-monitoring
7 judgments about specific instances, thus reducing the reporting of internal intrusions
8 (Lindsay et al., 2004; Oeberst & Blank, 2012). Given how follow-up open questions are
9 related to improved particularization and instance discrimination, they were
10 incorporated in the MMIF as prompts that are tailored to the interviewees' accounts
11 rather than as pre-set questions (*witness-compatible questioning*; Fisher & Geiselman,
12 1992; for a review see Oxburgh et al., 2010). Given the use of evidence-based cues,
13 more accurate information and fewer internal intrusions should be reported in the
14 MMIF and in the SGC-Timeline condition than in the Free Recall condition. However,
15 the use of follow-up prompts should further improve reporting and source monitoring in
16 the MMIF condition.

17 Similar to previous research on repeated events, some target activities were
18 manipulated to change from one instance to another, while some remained stable across
19 events (e.g., Brubacher et al., 2012; McNichol et al., 1999). To examine if the presence
20 of deviations at encoding facilitates recall and particularization, half of the participants
21 witnessed four events on separate occasions, where the third instance included two
22 deviations (one changed and one novel detail). We expected that participants in the
23 'deviation present' condition would recall more correct information both for the third
24 instance (targeted effect; e.g., Mac Lean et al., 2018) and across all events (general

1 effect; e.g., Connolly et al., 2016), compared to participants in the ‘deviation absent’
2 condition where no deviations were introduced.

3 **Method**

4 *Participants and Design*

5 A total of 150 participants (121 Females, Age: M = 21.26, SD = 5.21, Range 18-44
6 years) were randomly allocated to a 3 (Reporting format: Multi-Method Interviewing
7 Format (MMIF) vs. SGC-Timeline vs. Free recall) x 2 (Deviation: Present vs. Absent)
8 between-subjects design. Participants were recruited through the student participation
9 pool and advertisements circulated across campus and were granted course credit or a
10 £7 honorarium for participating. Overall, 164 participants were recruited but 14 did not
11 attend all the sessions and so were excluded from analyses. Dependent variables were
12 the number of correct details, correct gist and verbatim details, accuracy rates for all
13 types of reported details, and intrusion errors. Based on previous findings of the initial
14 testing of the timeline technique on the reporting of attributions of actions and
15 statements to people, we also included this dependent variable (i.e., number of reported
16 attributions and accuracy rate) to examine reporting across repeated events.

17 *Materials*

18 *Stimulus events.* Five stimulus events were scripted and filmed. Each event was
19 a short film, 4-5 minutes long, depicting a meeting between four perpetrators (three
20 males, one female) who plot a terrorist attack and then proceed to carry out the plan.
21 Each event was shot from a first-person perspective to facilitate the cover story that the
22 participant is an undercover agent acting as a group member. In each film, the leader
23 delivers information to the perpetrators about the target of the attack and assigns the

1 following roles to each member: one member will oversee the operation and provide the
2 detonator (a mobile phone) to another member who will plant the explosives; the third
3 member will act as a look out, and the participant will be the getaway driver. There is a
4 discussion about the explosives, how they are to be detonated and when. The meetings,
5 which constitute the first part of the event, take place indoors and were all shot in the
6 same location. In the second part of the event, the perpetrators are seen arriving at the
7 selected location, which differed in each film, and act according to their assigned roles
8 before they leave in a getaway car. The first part of the indoor conversations was highly
9 similar across events, although specific details of the content of the discussions varied.
10 The second part of the outdoor activities was similar in the overall structure, and on the
11 general level i.e., people involved, actions performed, but the location and direction of
12 people's movements varied (see Table 1 for all variable details across events; see Table
13 2 in Supplemental materials for a visual representation of the events).

14 *Insert Table 1*

15 To implement the deviation, an alternative version of one of the four events was
16 developed including two deviations: i) a 'role switch': one of the four perpetrators, who
17 always has the role of the lookout during the operation, was also in charge of the
18 meeting, while the perpetrator who always has the role of the leader simply attended the
19 meeting with the other members; and ii) a 'new character': when carrying out the plan,
20 the perpetrator in charge of planting the explosives gestured to the female overseeing
21 the operation to convey that there was a problem with the explosives. The female was
22 seen making a phone call, and a woman, who was not seen in the other events, briefly
23 appeared and handed her an envelope. This alternative version was only presented in the
24 Deviation Present condition and was always presented third to avoid primacy or recency
25 effects on recall. The presentation order of all the other events in both the Deviation

1 Present and Deviation Absent conditions was counterbalanced across participants to
2 avoid order effects. Analysis using a univariate analysis of variance showed that there
3 was no effect of the counterbalancing stream of 16 iterations used on the reporting of
4 correct details, $F(15, 149) = .92, p = .544, \eta^2 = .09$, nor on the reporting of correct gist
5 details, $F(15, 149) = 1.37, p = .170, \eta^2 = .13$.

6 *Timeline Reporting Format.* The two-level timeline format for reporting
7 repeated events consisted of: (i) a “Scoping” timeline (33 in. x 12 in.) which depicts a
8 horizontal line running at mid-point from one end of the card to the other to provide an
9 overview of all the experienced events; (ii) “Specific event” timelines of the same size
10 and layout (33 in. x 12 in.) which depicts a horizontal line running at mid-point from
11 one end to the other representing the temporal space for each event; (iii) Person
12 Description cards (5 in. x 3 in. white lined cards); (iv) Action cards (3 in. x 3 in.):
13 yellow cards (semi-adhesive strip on the back for easy removal and rearrangement on
14 the timeline); and (v) Statement cards (5 in. x 3 in.): blank, pink cards.

15 *Follow-up open-ended questions.* A protocol of open-ended in-depth questions
16 was composed to prompt additional information based on the initial account, in relation
17 to information gaps, and inconsistencies/clarifications (see Table 2).

18 *Insert Table 2*

19 **Procedure**

20 *Stimuli administration.* Participants visited the lab to witness four events on four
21 separate occasions over the span of seven days (minimum one day and maximum four
22 days between visits). On each occasion participants were instructed to imagine that they
23 were an undercover agent who had infiltrated a terrorist group. They were asked to pay
24 attention because they would later have to provide a report on the activities of the

1 group, which would be passed on to intelligence analysts. Participants witnessed the
2 events on a computer screen while wearing headphones. After witnessing the final
3 event, participants were invited to return after a seven-day delay ($M = 7.29$ days, $SD =$
4 0.53) to provide an account of the events. All participants provided their account
5 approximately two weeks after the first visit and one week after the last visit to the lab.

6 *Interview.* When participants returned to provide their account, they were either
7 provided with instructions for the Multi-Method Interviewing Format (MMIF), the SGC
8 - Timeline, or the Free recall format. All participants were reminded that they are in the
9 role of an undercover agent who infiltrated a terrorist group and that they are in
10 possession of valuable intelligence information about the activities of the group. When
11 asked to describe each event, all participants across conditions received the same
12 general instruction: ‘report all the details you remember about the events and the people
13 involved; report exactly what was said when possible’. Also, all participants were
14 instructed to not guess about things they could not remember.

15 In the MMIF condition, participants were instructed to begin by outlining all the
16 events in the order they witnessed them and to then focus on each instance. Participants
17 could use a card to label each event and place them on the scoping timeline. Next,
18 following Kontogianni et al. (2018), participants received the self-generated cue
19 instructions: ‘*Without thinking too hard, write down the first six things that you*
20 *remember seeing or thinking when witnessing each event. It doesn’t matter what these*
21 *things are. All that is important is that they immediately come to mind when thinking*
22 *back to each event. Please list them on a piece of paper. Think about each of the things*
23 *in your list one at a time and think about whether that memory helps you remember*
24 *other things that also happened in the event*’. Participants were instructed to list the
25 self-generated cues for each individual event, prior to using a timeline to describe each

1 event in detail. At that point, they were instructed to use the person description cards to
2 report descriptive details about each person, and the action and statement cards to report
3 action, sequence information and any statements they remembered. They were also
4 instructed to link the cards to show “who did/said what and when”.

5 After they finished providing their account in written form, participants were
6 asked follow-up open-ended questions. The interviewer asked three to four questions
7 per event, so that all participants across conditions were asked an equivalent number of
8 questions. Consistent with best practice interviewing guidance, the questions topics
9 were not pre-selected, instead participants were asked open questions based on what
10 they had reported in their account. For instance, if the participant had mentioned the
11 leader of the group, the interviewer asked “You mentioned there was a leader of the
12 group. Tell me more about this leader.” Or, if explosives had been mentioned in the
13 initial account, the interviewer asked, “Explain in more detail what you mean about this
14 part where they discussed the explosives”. This procedure allowed for interviewers to
15 maintain the same phrasing of questions but avoid using a scripted list of cued recall
16 questions not related to the witness’s initial account. Participants were not required to
17 answer all the questions and if they replied by saying “I don’t know” or “I don’t
18 remember”, the interviewer moved on to the next question. Finally, all participants were
19 asked if there was anything else they would like to report. During the questioning phase,
20 the participant’s written account remained in sight and the interviewer pointed to the
21 specific part to which the prompt referred when asking each question. The follow-up
22 questioning phase was audio and video-recorded, but the camera was only focused on
23 the participant’s written account. Participants in the SGC-Timeline condition followed
24 the exact same procedure in reporting their initial account in written form, but they were
25 not asked any follow-up questions. Participants in the Free recall condition were

1 instructed to outline the events prior to describing each in detail, but they did not
2 receive specific instructions on how to outline the events. After providing their account
3 in written form, they were asked follow-up questions according to the procedure
4 outlined above. At the end, participants were debriefed and compensated for their time.

5 *Coding*

6 Following Hope et al. (2019), a coding protocol was developed for the stimuli events.
7 Each detail reported was coded as a person, action, object or setting detail. Details were
8 coded as accurate² if they were present in each corresponding stimulus event and
9 described correctly. Particularisation refers to the reporting of (highly) specific details
10 as well as to the increased reporting of details that are specific to each instance of
11 repeated events (Brubacher & Roy, 2014; Powell et al. 2007). Use of the current coding
12 scheme allowed for particularisation to be reflected in our measure of correct recall.
13 Details that were specific to each instance (e.g., about the location and placement of
14 explosives; the equipment used) were reported and coded as following: ‘the female (1-
15 P) walked (1-A) into the building (1-S), into the lobby (1-S) and to the right (1-S)’;
16 ‘carrying (1-A) a black (1-O) backpack (1-O) with red markings (1-O) and a mobile
17 phone (1-O).’ Each instance took place in a different location and setting; therefore, this
18 example refers to the only one occasion where the female perpetrator placed explosives
19 in a building lobby. Details that were vague or subjective (e.g., “he was young”, “he
20 looked satisfied”) were not scored for accuracy. As the events included a conversation
21 among the perpetrators, interviews were also coded for gist and verbatim statements,

² We acknowledge the potential for two definitions of reported accuracy (narrow and broad) as described in the meta-analysis by Woiwod et al. (2019). Given the framing on particularization and the hypotheses of the current research, we formulated our instructions and developed our coding scheme to deliberately focus on the amount of correct information about each instance of the repeated events.

1 based on the script that was developed for the stimulus events. Gist details reflected the
2 overall meaning of what was discussed (e.g., “the leader told everyone what their role
3 was”) and were scored as one point for each correct gist unit (i.e., correct extraction of
4 the conversation that was not reported verbatim) and one point for each incorrect gist
5 unit (i.e., incorrect extraction). If the gist statement was reported in a vague manner, it
6 was not scored for accuracy (e.g., “they talked about doing something”). Verbatim
7 details reflected the precise language used in the original stimulus. Verbatim units were
8 scored as correct for every three verbatim words reported correctly and as incorrect
9 when two or fewer words corresponded to the script. In the initial account, additional
10 coding was conducted for the accuracy of attributions of both actions and statements to
11 a person. Attribution details were scored as correct when an action or statement was
12 correctly attributed to a specific actor (e.g., Female handed the detonator over).

13 All the accounts were coded for internal intrusions (i.e., source monitoring
14 errors) by noting the type of the reported detail (person, action, object, setting, gist,
15 verbatim, location, target, time; see Table 1) and the source of the stimulus event where
16 it was witnessed. For example, if a participant reported that the “target in the Spinnaker
17 tower was a local activist”, that would be coded as an intrusion as this was the target in
18 the event at Victoria park. Therefore, if the event at Victoria park was witnessed third,
19 this would be scored under intrusions as “1-Target Event-3”. The same coding scheme
20 was used to code the responses to the follow-up questions, apart from coding for the
21 attribution details which were only relevant to the initial reporting phase.

22 Coding of the interviews was mostly conducted by the lead researcher and partly
23 by two research assistants. Twenty-four interviews (i.e., 15% of all interviews) were
24 randomly selected and independently scored by the second author, who was blind to
25 experimental conditions for Deviation (to some extent also for Format; i.e., between the

1 MMIF and SGC-Timeline conditions which used the same reporting format). Inter-rater
2 reliability, which was high across coding categories, $ICC = .97$, 95% CI [.967, .974],
3 was computed based on the mean value of two raters, using an absolute agreement
4 definition and a two-way mixed effects model (McGraw & Wong, 1996).

5 ***Statistical Analyses***

6 To examine the extent to which the independent variables predicted correct reporting,
7 the dependent variables were analysed using linear mixed models (LMMs) with fixed
8 effects of reporting format (categorical: MMIF vs. SGC-Timeline vs. Free recall) and
9 deviation (categorical: Absent vs Present), and random intercepts for events nested
10 within participants³ (Finch et al., 2014). We arrived at this model by comparing: i) a
11 baseline model with fixed effects only; and ii) a model with fixed effects and random
12 intercepts for events nested within participants. We found that the second model was a
13 better fit for the data, by conducting a likelihood ratio test (LRT, function *anova*)
14 comparing the log-likelihoods of both models (for all the statistical comparisons of the
15 models see Supplemental Materials). The model included two-way interactions between
16 deviation and reporting format. We used simple contrasts to code the reporting format
17 and the deviation. Specifically, for format, contrasts compared reporting with Free
18 recall (reference level) to reporting with SGC-Timeline, and Free recall to MMIF. A
19 separate model was used with a contrast comparing MMIF to SGC-Timeline. For
20 deviation, the contrasts compared Absent vs Present. The reported coefficients (*b* value
21 estimates) show the degree to which the dependent variable changed relative to the
22 reference level, while the 95% CI represent the plausible range of the value of the

³ The only exception was for the analysis of the number of intrusions reported across events, where random intercepts for events were nested within type of intrusions within participants.

1 regression coefficients (Cumming, 2013). With respect to accuracy rates, we were
2 interested in the overall accuracy reported across events rather than within each
3 instance, and thus analysis was not conducted with LMMs, but with factorial ANOVAs.
4 Accuracy rates were calculated by dividing the number of correct details reported by the
5 total number of details (correct and incorrect) reported.

6 The analyses were run in R version 3.5.0 (R Core Team, 2017) using the `lme`
7 function from the `nlme` package (Pinheiro et al., 2017). Datasets, and R markdown
8 scripts are available on the Open Science Framework website (<https://osf.io/4mcsa/>).

9 **Results**

10 We present the results of the analysis on the number of correct details, related accuracy
11 rate, and number of gist details for total reporting⁴ first, and then examine the initial
12 reports and the reporting of internal intrusions. In the initial reporting phase, we also
13 present secondary analyses with respect to the number of correct attributions of actions
14 and statements to people, and accuracy rate of attribution details. In the interest of
15 parsimony, results for verbatim details and accuracy rates for both gist and verbatim
16 details compared for initial and total reporting, are presented in Supplemental Materials,
17 as no significant results emerged across conditions.

18 ***Total interview output***

19 Participants in the MMIF condition were asked a similar number of follow-up questions
20 ($M = 13.78$, $SD = 1.58$) to the participants in the Free recall condition ($M = 13.16$, $SD =$

⁴ To compare the differences between the Reporting formats as a whole regarding the information reported per instance, we built a model that examined total reporting rather than reporting in each interview phase. This approach was in line with the current research aims and hypotheses. The information gain in the follow-up phase is also evident with the use of this model (see Figures 1 and 3).

1 1.60). An independent *t*-test analysis showed that there was no statistically significant
2 difference in the number of follow-up questions asked between groups, $t(98) = 1.95$, $p =$
3 $.054$.

4 Reporting Format had a significant main effect on the reporting of total correct
5 details, with more correct details being reported in the MMIF condition, $b = 7.48$, 95%
6 $CI [0.98, 13.98]$, $t(144) = 4.42$, $p < .001$, and in the SGC-Timeline condition, $b = 5.48$,
7 95% $CI [-1.02, 11.98]$, $t(144) = 3.24$, $p = .001$ than in the Free Recall condition.

8 Reporting of total correct details did not differ between the SGC-Timeline and the
9 MMIF condition, $b = 2.00$, 95% $CI [-4.50, 8.50]$, $t(144) = 1.18$, $p = .240$. Neither
10 Deviation, $b = 0.95$, 95% $CI [-6.26, 4.36]$, $t(144) = 0.69$, $p = .491$, nor the interaction
11 between Deviation and Reporting format were significant: i) MMIF vs Free Recall, $b =$
12 0.46 , $t(144) = 0.14$, $p = .889$; ii) MMIF vs SGC-Timeline, $b = 0.18$, $t(144) = 0.05$, $p =$
13 $.960$; iii) SGC-Timeline vs Free Recall, $b = 0.28$, $t(144) = 0.08$, $p = .936$. Therefore, the
14 use of different formats only affected the reporting of correct details (see Figure 1, for
15 results at initial and follow-up reporting).

16 *Insert Figure 1*

17 A separate model was built to examine if there is a targeted effect of Deviation
18 on the reporting of correct details for the instance that contained the deviations.
19 Deviation did not have a significant effect on the reporting of correct details, $b = 2.53$,
20 95% $CI [-1.10, 6.17]$, $t(144) = 1.38$, $p = .170$. Therefore, the presence of deviations did
21 not affect the reporting of details overall (general effect on recall) or for the specific
22 instance (target effect on recall). Figure 2 shows the reporting of correct details across
23 events within deviation conditions (total reporting).

24 *Insert Figure 2*

1 A factorial ANOVA showed that neither Reporting Format, $F(1, 96) = 0.21, p =$
2 $.650, \omega^2 = .009$, nor Deviation, $F(1, 96) = 1.04, p = .311, \omega^2 = .010$, had a significant
3 main effect on the overall accuracy of accounts. The interaction between Deviation and
4 Reporting format was not significant, $F(1, 96) = 0.93, p = .337, \omega^2 = .000$.

5 A repeated-measures analysis of variance showed that the accuracy rate of the
6 reported information in the follow-up questioning phase was significantly lower than
7 the accuracy rate in the initial reporting phase, $F(1, 98) = 32.54, p < .001, \omega^2 = .245$.
8 However, accuracy in the follow-up questioning phase was not affected by the format
9 participants used to provide their initial report, $F(1, 98) = 1.38, p = .243, \omega^2 = .003$.
10 Means and standard deviations for accuracy rates can be found in Table 3.

11 *Insert Table 3*

12 Reporting format had a significant effect on the reporting of correct gist details,
13 with more correct gist details reported in the MMIF than in both the Free recall, $b =$
14 $0.78, 95\% CI [0.09, 1.46], t(144) = 2.24, p = .026$, and the SGC-Timeline condition, $b =$
15 $1.47, 95\% CI [0.14, 2.79], t(144) = 4.24, p < .001$. More correct gist details were
16 reported in the Free recall than in the SGC-Timeline condition, $b = 0.69, 95\% CI [-1.37,$
17 $-0.01], t(144) = 2.00, p = .047$. Deviation did not have an effect on the reporting of
18 correct gist details $b = 0.06, 95\% CI [-0.50, 0.32], t(144) = 0.21, p = .832$. The
19 interaction between Deviation and Reporting format was not significant: i) MMIF vs
20 Free Recall, $b = 0.47, t(144) = 0.68, p = .498$; ii) MMIF vs SGC-Timeline, $b = 0.79,$
21 $t(144) = 1.14, p = .256$; iii) SGC-Timeline vs Free Recall, $b = .32, t(144) = 0.46, p =$
22 $.646$.

23 Figure 3 presents the reporting of correct gist details as a function of Reporting
24 format within Deviation conditions. Figure 4 shows the reporting of correct gist details
25 across events within Deviation conditions (total reporting).

1 *Insert Figure 3*

2 *Insert Figure 4*

3 ***Initial reporting phase***

4 Reporting format had a significant main effect on the number of correct details reported,
5 with more correct details reported initially in the MMIF condition, $b = 9.04$, 95% *CI*
6 [2.89, 15.19], $t(144) = 5.65$, $p < .001$, and in the SGC-Timeline condition, $b = 12.47$,
7 95% *CI* [6.32, 18.62], $t(144) = 7.80$, $p < .001$ (cf. Free Recall). More correct details
8 were reported in the SGC-Timeline than in the MMIF condition, $b = 3.43$, 95% *CI* [-
9 2.72, 9.58], $t(144) = 2.14$, $p = .034$. Neither Deviation, $b = 1.03$, 95% *CI* [-6.05, 3.98],
10 $t(144) = 0.79$, $p = .431$, nor an interaction between Deviation and Reporting format
11 were significant: i) MMIF vs Free Recall, $b = 1.50$, $t(144) = 0.47$, $p = .640$; ii) MMIF vs
12 SGC-Timeline, $b = 0.58$, $t(144) = 0.18$, $p = .857$; iii) SGC-Timeline vs Free Recall, $b =$
13 0.92, $t(144) = 0.29$, $p = .772$.

14 A factorial ANOVA showed a significant main effect of Reporting format on the
15 accuracy rate of reported details, $F(2,144) = 3.43$, $p = .035$, $\omega^2 = .032$. Bonferroni-
16 adjusted post-hoc pairwise comparisons showed that there was a significant difference
17 ($p = .035$) between the MMIF ($M = 0.87$, $SD = 0.06$) and the Free recall condition ($M =$
18 0.83, $SD = 0.08$), but not between the SGC-Timeline and Free recall conditions ($p =$
19 .228). There was not a significant difference between the MMIF and SGC-Timeline
20 conditions ($p = 1.00$). There was no significant main effect of Deviation on the
21 accuracy rate $F(1,144) = 0.14$, $p = .709$, $\omega^2 = -.006$, and the interaction between
22 Reporting format and Deviation was not significant $F(2, 144) = 2.47$, $p = .088$, $\omega^2 =$
23 .020.

1 Reporting format had a significant main effect on the reporting of gist details,
2 with more correct gist details reported initially in the MMIF than the Free recall format,
3 $b = 0.67$, 95% *CI* [0.02, 1.32], $t(144) = 2.03$, $p = .044$. Reporting of correct gist details
4 did not differ between the MMIF and SGC-Timeline conditions, $b = 0.21$, 95% *CI* [-
5 0.86, 0.44], $t(144) = 0.64$, $p = .523$, or between the SGC-Timeline and Free recall
6 conditions, $b = 0.46$, 95% *CI* [-0.19, 1.11], $t(144) = 1.39$, $p = .167$. Deviation did not
7 have a significant main effect on the reporting of gist details ($p = .758$) and the
8 interaction between Reporting format and Deviation was not significant: i) MMIF vs
9 Free Recall, $b = 0.22$, $t(144) = 0.33$, $p = .742$; ii) MMIF vs SGC-Timeline, $b = 0.70$,
10 $t(144) = 1.06$, $p = .291$; iii) SGC-Timeline vs Free Recall, $b = 0.48$, $t(144) = 0.73$, $p =$
11 $.467$.

12 Reporting format had a significant main effect on the reporting of correct
13 attributions of actions and statements to persons, with more attributions reported in the
14 MMIF, $b = 1.58$, 95% *CI* [0.72, 2.43], $t(144) = 3.63$, $p < .001$, and the SGC-Timeline
15 conditions, $b = 2.38$, 95% *CI* [1.52, 3.24], $t(144) = 5.49$, $p < .001$, than in the Free recall
16 condition. Reporting was not significantly different in SGC-Timeline and MMIF
17 conditions, $b = 0.81$, 95% *CI* [-0.05, 1.66], $t(144) = 1.86$, $p = .065$. Deviation did not
18 have a significant main effect on the correct reporting of attributions, $b = 0.23$, 95% *CI*
19 [-0.93, 0.47], $t(144) = 0.66$, $p = .510$, and the interaction between Deviation and
20 Reporting format was not significant, i) MMIF vs Free Recall, $b = 0.07$, $t(144) = 0.08$, p
21 $= .936$; ii) MMIF vs SGC-Timeline, $b = 0.25$, $t(144) = 0.29$, $p = .773$; iii) SGC-Timeline
22 vs Free Recall, $b = 0.18$, $t(144) = 0.21$, $p = .834$. Figure 5 shows the number of correct
23 attributions made as a function of Reporting format within Deviation conditions.

24 *Insert Figure 5*

1 A factorial ANOVA showed that Reporting format had a main effect on the
2 accuracy rate of the reported attribution details, $F(2,144) = 4.72, p = .010, \omega^2 = .048$.
3 Deviation did not have a significant effect, $F(1,144) = 0.18, p = .672, \omega^2 = -.315$, and
4 the interaction between Reporting format and Deviation was not significant, $F(2,144) =$
5 $0.11, p = .897, \omega^2 = -.161$. Levene's test was significant ($p = .010$). Bonferroni-adjusted
6 post-hoc pairwise comparisons showed that there were significantly higher accuracy
7 rates of reported attributions in the MMIF ($M = 0.73, SD = 0.16$) than in the Free recall
8 condition ($M = 0.61, SD = 0.29$), $p = .023$, and in the SGC-Timeline ($M = 0.73, SD =$
9 0.19) than in the Free recall condition ($M = 0.61, SD = 0.29$), $p = .028$, with no
10 difference between the MMIF and SGC-Timeline conditions ($p = 1.00$).

11 ***Internal intrusions across instances***

12 Table 4 shows the distribution of the mean number of internal intrusions in each event
13 across conditions.

14 *Insert Table 4*

15 Neither Deviation, $b = 0.01, 95\% CI [-0.06, 0.05], t(144) = 0.44, p = .661$, nor
16 Reporting format had a significant main effect on the reporting of intrusions: i) MMIF
17 vs Free Recall, $b = 0.02, 95\% CI [-0.04, 0.08], t(144) = 1.23, p = .221$; ii) MMIF vs
18 SGC-Timeline, $b = 0.02, 95\% CI [-0.04, 0.08], t(144) = 1.23, p = .221$; iii) SGC-
19 Timeline vs Free Recall, $b = 0.00, 95\% CI [-0.06, 0.06], t(144) = 0.00, p = 1.00$. The
20 interaction between Deviation and Reporting format was significant for i) SGC-
21 Timeline vs Free Recall, $b = 0.08, t(144) = 2.31, p = .022$ but not for: ii) MMIF vs Free
22 Recall, $b = 0.02, t(144) = 0.54, p = .590$; iii) MMIF vs SGC-Timeline, $b = 0.06, t(144) =$
23 $1.77, p = .079$.

1 **Discussion**

2 The current study examined the effectiveness of the timeline technique used in
3 combination with self-generated cues and follow-up open-ended questions, as a multi-
4 method interviewing format, for facilitating recall and particularization of repeated
5 events (i.e., correct information specific to each instance). The findings show that
6 participants reported overall more correct information about specific instances when
7 using the MMIF than a free recall format followed by open-ended questions, without a
8 cost to accuracy. Findings for the MMIF and SGC-timeline conditions were comparable
9 with respect to the reporting of information about the attacks, but as more information
10 was reported in the MMIF condition about the conversations around planning the attack,
11 reporting was overall higher in the MMIF (cf. SGC-Timeline) condition.⁵

12 Although the responses to follow-up questions were not as accurate as the
13 information provided in the initial accounts, previous research suggests that an increase
14 in the reporting of both correct details and errors is likely when output increases overall
15 (Memon et al., 2010; Roberts & Higham, 2002; see also, Kontogianni et al., 2020 for a
16 replication of this result). Importantly, the use of follow-up questions about each
17 instance facilitated the reporting of more episodic information, and accuracy rates were
18 similar across reporting format conditions. For instance, the use of follow-up questions
19 allowed participants to elaborate on some details, thus facilitating further reporting, but
20 more effectively so when participants had already provided a detailed account using
21 effective mnemonics as shown in the MMIF condition.

⁵ Although there was a main effect of Reporting format on the time of reporting, $F(2, 142) = 76.96, p < .001$, Bonferroni post-hoc comparisons show that the reporting time in the Free recall condition was significantly lower than in the MMIF and SGC-Timeline conditions ($p < .001$), but that reporting time did not differ between the MMIF and SGC-Timeline conditions ($p = 1.00$). Therefore, the use of follow-up open questions in the MMIF condition did not result in a significantly longer interview, compared to the other condition where the timeline technique was used.

1 A closer analysis of the initial reporting phase shows that the initial accounts
2 were more detailed about each instance when participants used the self-generated cues
3 in conjunction with the timeline technique rather than a free recall format. Based on
4 previous research comparing the separate individual components of the timeline
5 technique (i.e., full timeline, record cards, temporal context instructions; Experiment 2;
6 Hope et al., 2013) to a free recall format, it is likely that the mnemonic benefits inherent
7 in the timeline format and the use of separate timelines per instance in the current study
8 facilitated the retrieval of specific details (see also Hope et al., 2019). It is possible that
9 particularization was further facilitated by the use of the self-generated cues, which
10 were used to prompt the retrieval of the salient details prior to describing each instance
11 in detail (Brubacher et al., 2011). However, based on the current method we cannot be
12 certain about the contribution of the use of the self-generated cues to overall
13 performance. Consistent with previous findings, participants who used the timeline
14 technique also reported more correct attributions of actions and statements to
15 perpetrators than those who used a free recall format (Hope et al., 2013; 2019).
16 Regarding the use of free recall, and potential concerns that any additional intervention
17 would facilitate retrieval, we note that this comparison allowed the examination of self-
18 administered reporting and matching of the general instructions across conditions. Also,
19 participants in the free recall condition still had the opportunity to report more
20 information in response to follow-up questions compared to participants in the SGC-
21 Timeline condition. The current results provide further evidence of the usefulness of the
22 timeline technique in conjunction with self-generated cues for multi-actor events in the
23 reporting of both single and repeated experiences.

24 The increased reporting and particularization of instances in both conditions
25 where the timeline technique and self-generated cues were used (cf. free recall)

1 occurred without an increased cost in the reporting of internal intrusions. Given
2 evidence that the use of open in-depth questions per instance elicit more specific details
3 and facilitate discrimination between instances (see also, Brubacher et al., 2012), fewer
4 intrusions were expected in the MMIF relative to the SGC-Timeline and to the free
5 recall conditions. However, the mean number of intrusions was low across conditions.
6 Notably, the number of intrusions varied by detail type. Participants most commonly
7 confused the equipment used across events and the perpetrators who alternated roles in
8 planting the explosives, with the latter being a particularly pertinent detail for
9 interviewers who are interested in knowing ‘what happened when’ or ‘who did what
10 when’ (Roberts, 2003). Considering that internal intrusions are unlikely to be
11 completely avoided, the current findings suggest that corroboration is necessary for
12 details that interviewees have difficulty attributing to an instance. Inquiring about which
13 instance interviewees remember the best, and why, could help us better understand how
14 interviewees encode and process information about repeated events (Danby et al.,
15 2017). Given the current null finding, further research is necessary to determine the
16 precise impact of the use of the MMIF components on source monitoring.

17 Contrary to our hypotheses, the presence of deviations had no effect on either
18 recall or on particularization of specific instances. Accurate reporting was improved
19 when memory-enhancing techniques were used (cf. free recall format), yet recall did not
20 further benefit from the presence of deviations. The current results are inconsistent with
21 previous research, which shows that the presence of deviations facilitates recall for all
22 the witnessed events and for the targeted instance itself (Connolly et al., 2016; MacLean
23 et al., 2018; although see Rubinova et al., 2020). Also, witnessing an instance that
24 deviated from the script did not facilitate source monitoring as participants reported a
25 similar number of internal intrusions in both deviation conditions. One explanation for

1 our results may be that the current deviations were not sufficiently salient⁶ to impact
2 recall. To this end, we could have included a manipulation check to directly inquire if
3 participants encoded the deviations (e.g., MacLean et al., 2018), as it is difficult to reach
4 any conclusions based on what participants reported alone.

5 Research suggests that the effect of deviations on recall may depend on whether
6 the deviation has any consequence to the sequence of events (i.e., *continuous vs discrete*
7 *deviations*; Connolly et al., 2016; *obstacles, errors, and distractions*; Schank &
8 Abelson, 1977). Although the comparison between specific types of deviations was
9 beyond the focus of our study, exploring how deviations are implemented across studies
10 and their differential effect on recall might be relevant for future research. Notably, in
11 recent studies showing improved recall for deviations, both children and adult
12 participants witnessed the repeated events over one or two days (e.g., Connolly et al.,
13 2016; MacLean et al., 2018; although see Brubacher et al., 2012) and usually with recall
14 occurring after shorter intervals than in the current study. It is possible, then, that after
15 the one-week interval, our participants could not access the deviation details. Further
16 research on the effects of deviations on memory is needed to determine if the lack of an
17 observed effect depends more on the salience of the deviations, or the delay until
18 reporting. An interaction between the two factors is also conceivable.

19 Lastly, it is possible that a larger sample size may be required to detect an effect
20 of the presence of deviations on recall. Key reasons, identified before conducting the
21 research, prevented us from conducting an a priori analysis. First, in the repeated events

⁶ We can report that although the ‘changed’ detail (i.e., a different person leading the group meeting on one occasion) was reported by participants, the ‘new’ detail (i.e., a new person briefly entering the scene in the park, passing by one of the perpetrators to hand her an envelope) was not reported. Possibly, the ‘new’ detail was not conspicuous enough in a busy scene; or maybe participants used the existing script to notice the ‘changed’ compared to the ‘new’ detail, since the script for the former was likely stronger than the latter, which occurred during the second and more variable part of the events (*schema-confirmation-deployment hypothesis*; Farrar & Goodman, 1992).

1 literature, recall is often measured in a different way than in the current experiment
2 (e.g., reporting of pre-determined fixed vs variable details). Second, power calculations
3 for Linear Mixed Models require the input of parameter estimates that are often not
4 reported in the literature and/or are tied to the stimuli used (Kumle et al., 2020). These
5 estimates were not available to the current research.

6 The current findings suggest that self-generated cues, the timeline technique,
7 and follow-up open-ended questions can be used together to elicit detailed and accurate
8 information for repeated events. With respect to the manipulation of deviations, our
9 results suggest that further research should investigate how deviations affect (or not)
10 delayed recall to improve reporting and discrimination between instances. The current
11 research contributes to the development of an adaptive information gathering ‘toolbox’
12 of techniques that can be used flexibly in applied settings. However, further research is
13 needed to address the challenges that both interviewers and interviewees face when
14 investigating repeated events, including the implications of interference between
15 instances for reporting.

16 **Acknowledgements**

17 This work was funded by the Centre of Research and Evidence on Security Threats
18 (ESRC Award: ES/N009614/1) and the University of Portsmouth as a postgraduate
19 bursary for the doctoral work of the first author.

20 **References**

- 21 Abelson, R. P. (1981). Psychological status of the script concept. *American*
22 *Psychologist*, 36(7), 715-729. <http://dx.doi.org/10.1037/0003-066X.36.7.715>
- 23 Ahn, W. K., Brewer, W. F., & Mooney, R. J. (1992). Schema acquisition from a single
24 example. *Journal of Experimental Psychology: Learning, Memory, and*
25 *Cognition*, 18(2), 391-412. <http://dx.doi.org/10.1037/0278-7393.18.2.391>

- 1 Anderson, J. R. (1983). A spreading activation theory of memory. *Journal of Verbal*
2 *Learning and Verbal Behavior*, 22(3), 261–295. [doi:10.1016/S0022-](https://doi.org/10.1016/S0022-5371(83)90201-3)
3 [5371\(83\)90201-3](https://doi.org/10.1016/S0022-5371(83)90201-3)
- 4 Anderson, S. J., & Conway, M. A. (1993). Investigating the structure of
5 autobiographical memories. *Journal of Experimental Psychology: Learning,*
6 *Memory, and Cognition*, 19(5), 1178–1196. [https://doi.org/10.1037/0278-](https://doi.org/10.1037/0278-7393.19.5.1178)
7 [7393.19.5.1178](https://doi.org/10.1037/0278-7393.19.5.1178)
- 8 Belli, R. F. (1998). The structure of autobiographical memory and the event history
9 calendar: Potential improvements in the quality of retrospective reports in
10 surveys. *Memory*, 6(4), 383-406. [doi:10.1080/741942610](https://doi.org/10.1080/741942610)
- 11 Brainerd, C. J., & Reyna, V. F. (2004). Fuzzy-trace theory and memory
12 development. *Developmental Review*, 24(4), 396-439.
13 <https://doi.org/10.1016/j.dr.2004.08.005>
- 14 Brainerd, C. J., & Reyna, V. F. (2001). *Fuzzy-trace theory: Dual processes in memory,*
15 *reasoning, and cognitive neuroscience.* In H. W. Reese & R. Kail
16 (Eds.), *Advances in child development and behavior* (p. 41–100). Academic
17 Press.
- 18 Brainerd, C. J., & Reyna, V. F. (1990). Gist is the grist: Fuzzy-trace theory and the new
19 intuitionism. *Developmental Review*, 10(1), 3-47. [https://doi.org/10.1016/0273-](https://doi.org/10.1016/0273-2297(90)90003-M)
20 [2297\(90\)90003-M](https://doi.org/10.1016/0273-2297(90)90003-M)
- 21 Brewer, W. F., & Nakamura, G. V. (1984). The nature and functions of schemas. *Center*
22 *for the Study of Reading Technical Report; no. 325.* Illinois, USA: University of

- 1 Illinois at Urbana-Champaign, Center for the Study of Reading. Bolt Beranek
2 and Newman, Inc. Retrieved from: <http://hdl.handle.net/2142/17542>
- 3 Brubacher, S. P., Earhart, B., Roberts, K. P., & Powell, M. B. (2018). Effects of label
4 training and recall order on children's reports of a repeated event. *Applied*
5 *Cognitive Psychology*, 32(5), 600-609. <https://doi.org/10.1002/acp.3440>
- 6 Brubacher, S. P., Glisic, U., Roberts, K. P., & Powell, M. (2011). Children's ability to
7 recall unique aspects of one occurrence of a repeated event. *Applied Cognitive*
8 *Psychology*, 25(3), 351-358. <https://doi.org/10.1002/acp.1696>
- 9 Brubacher, S. P., Powell, M. B., & Roberts, K. P. (2014). Recommendations for
10 interviewing children about repeated experiences. *Psychology, Public Policy,*
11 *and Law*, 20(3), 325-335. <http://dx.doi.org/10.1037/law0000011>
- 12 Brubacher, S. P., Roberts, K. P., & Powell, M. (2012). Retrieval of episodic versus
13 generic information: Does the order of recall affect the amount and accuracy of
14 details reported by children about repeated events?. *Developmental*
15 *Psychology*, 48(1), 111-122. <https://doi.org/10.1037/a0025864>
- 16 Brubacher, S. P., & La Rooy, D. (2014). Witness recall across repeated interviews in a
17 case of repeated abuse. *Child abuse & neglect*, 38(2), 202-211.
18 <https://doi.org/10.1016/j.chiabu.2013.06.010>
- 19 Cohen, G., & Java, R. (1995). Memory for medical history: Accuracy of recall. *Applied*
20 *Cognitive Psychology*, 9(4), 273-288. <https://doi.org/10.1002/acp.2350090402>
- 21 Connolly, D. A., & Gordon, H. M. (2014). Can order of general and specific memory
22 prompts help children to recall an instance of a repeated event that was different

- 1 from the others?. *Psychology, Crime & Law*, 20(9), 852-864.
2 <https://doi.org/10.1080/1068316X.2014.885969>
- 3 Connolly, D. A., Gordon, H. M., Woiwod, D. M., & Price, H. L. (2016). What children
4 recall about a repeated event when one instance is different from the others.
5 *Developmental Psychology*, 52(7), 1038-1051.
6 <https://doi.org/10.1037/dev0000137>
- 7 Conway, M. A., & Pleydell-Pearce, C. W. (2000). The construction of autobiographical
8 memories in the self-memory system. *Psychological review*, 107(2), 261-
9 288. <https://doi.org/10.1037/0033-295X.107.2.261>
- 10 Cumming, G. (2013). *Understanding the new statistics: Effect sizes, confidence*
11 *intervals, and meta-analysis*. Routledge, USA.
- 12 Danby, M. C., Brubacher, S. P., Sharman, S. J., Powell, M. B., & Roberts, K. P. (2017).
13 Children's reasoning about which episode of a repeated event is best
14 remembered. *Applied Cognitive Psychology*, 31(1), 99-108.
15 <https://doi.org/10.1002/acp.3306>
- 16 Davidson, D. (2006). Memory for bizarre and other unusual events: Evidence from
17 script research. In R. Hunt & J. B. Worthen (Eds.), *Distinctiveness and memory*
18 (pp. 157–179). Oxford University.
- 19 Farrar, M. J., & Boyer-Pennington, M. E. (1999). Remembering specific episodes of a
20 scripted event. *Journal of Experimental Child Psychology*, 73(4), 266-288.
21 <https://doi.org/10.1006/jecp.1999.2507>

- 1 Finch, W. H., Bolin, J. E., & Kelley, K. (2014). *Multilevel modeling using R*. CRC
2 Press.
- 3 Fisher, R., & Geiselman, R. (1992). *Memory-enhancing techniques for investigative*
4 *Interviewing: The cognitive interview*. Charles C. Thomas.
- 5 Goh, W. D., & Lu, S. H. (2012). Testing the myth of the encoding–retrieval
6 match. *Memory & cognition*, 40(1), 28-39. [https://doi.org/10.3758/s13421-011-](https://doi.org/10.3758/s13421-011-0133-9)
7 [0133-9](https://doi.org/10.3758/s13421-011-0133-9)
- 8 Graesser, A. C., Woll, S. B., Kowalski, D. J., & Smith, D. A. (1980). Memory for
9 typical and atypical actions in scripted activities. *Journal of Experimental*
10 *Psychology: Human Learning and Memory*, 6(5), 503.
- 11 Hope, L., Gabbert, F., Kinninger, M., Kontogianni, F., Bracey, A., & Hanger, A. (2019).
12 Who said what and when? A timeline approach to eliciting information and
13 intelligence about conversations, plots and plans. *Law and Human Behavior*;
14 43(3), 263-277. <http://dx.doi.org/10.1037/lhb0000329>
- 15 Hope, L., Mullis, R., & Gabbert, F. (2013). Who? What? When? Using a timeline
16 technique to facilitate recall of a complex event. *Journal of Applied Research in*
17 *Memory and Cognition*, 2(1), 20-24. [doi:10.1016/j.jarmac.2013.01.002](https://doi.org/10.1016/j.jarmac.2013.01.002)
- 18 Hudson, J. A., Fivush, R., & Kuebli, J. (1992). Scripts and episodes: The development
19 of event memory. *Applied Cognitive Psychology*, 6(6), 483-505.
20 10.1002/acp.2350060604
- 21 Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring.
22 *Psychological Bulletin*, 114(1), 3-28.

- 1 Kontogianni, F., Hope, L., Taylor, P. J., Vrij, A., & Gabbert, F. (2018). The benefits of a
2 self-generated cue mnemonic for timeline interviewing. *Journal of Applied*
3 *Research in Memory and Cognition*. 7(3), 454-
4 461. [10.1016/j.jarmac.2018.03.006](https://doi.org/10.1016/j.jarmac.2018.03.006)
- 5 Kontogianni, F., Hope, L., Taylor, P. J., Vrij, A., & Gabbert, F. (2020). “Tell me more
6 about this...”: An examination of the efficacy of follow-up open questions
7 following an initial account. *Applied Cognitive Psychology*.
8 <https://doi.org/10.1002/acp.3675>
- 9 Kumle, L., Vo, M. L. H., & Draschkow, D. (2020). *Estimating power in (generalized)*
10 *linear mixed models: an open introduction and tutorial in R*. PsyArXiv,
11 <https://doi.org/10.31234/osf.io/vxfbh>
- 12 Leins, D. A., Fisher, R. P., Pludwinski, L., Rivard, J., & Robertson, B. (2014). Interview
13 protocols to facilitate human intelligence sources' recollections of
14 meetings. *Applied Cognitive Psychology*, 28(6), 926-935.
15 <https://doi.org/10.1002/acp.3041>
- 16 Lindsay, D. S., (2014). Memory Source monitoring applied. In Perfect, T. J. & Lindsay,
17 D. S (Eds.), *The SAGE Handbook of Applied Memory* (59-75). SAGE
18 Publications Inc.
- 19 Lindsay, D. S., & Johnson, M. K. (1989). The eyewitness suggestibility effect and
20 memory for source. *Memory & Cognition*, 17(3), 349-358.
21 <https://doi.org/10.3758/BF03198473>
- 22 MacLean, C. L., Coburn, P. I., Chong, K., & Connolly, D. L. (2018). Breaking script:
23 Deviations and postevent information in adult memory for a repeated

- 1 event. *Applied Cognitive Psychology*, 32(4), 474-486.
2 <https://doi.org/10.1002/acp.3421>
- 3 Mäntylä, T., & Nilsson, L.-G. (1988). Cue distinctiveness and forgetting: Effectiveness
4 of self-generated retrieval cues in delayed recall. *Journal of Experimental*
5 *Psychology: Learning, Memory, and Cognition*, 14(3), 502–
6 509. <https://doi.org/10.1037/0278-7393.14.3.502>
- 7 McGraw, K. O., & Wong, S. P. (1996). "Forming inferences about some intraclass
8 correlations coefficients": Correction. *Psychological Methods*, 1(4),
9 390. <https://doi.org/10.1037/1082-989X.1.4.390>
- 10 McNichol, S., Shute, R., & Tucker, A. (1999). Children's eyewitness memory for a
11 repeated event. *Child Abuse & Neglect*, 23(11), 1127-1139.
12 [https://doi.org/10.1016/S0145-2134\(99\)00084-8](https://doi.org/10.1016/S0145-2134(99)00084-8)
- 13 Means, B. & Loftus, E. (1991). When personal history repeats itself: Decomposing
14 memories for recurring events. *Applied Cognitive Psychology*, 5(4), 297–318.
15 <https://doi.org/10.1002/acp.2350050402>
- 16 Memon, A., Meissner, C. A., & Fraser, J. (2010). The Cognitive Interview: A meta-
17 analytic review and study space analysis of the past 25 years. *Psychology, Public*
18 *Policy, and Law*, 16(4), 340-372. <https://doi.org/10.1037/a0020518>
- 19 Nairne, J. S. (2002). The myth of the encoding-retrieval match. *Memory*, 10(5-6), 389-
20 395. [doi:10.1080/09658210244000216](https://doi.org/10.1080/09658210244000216)

- 1 Oeberst, A., & Blank, H. (2012). Undoing suggestive influence on memory: The
2 reversibility of the eyewitness misinformation effect. *Cognition*, 125(2), 141-
3 159. <https://doi.org/10.1016/j.cognition.2012.07.009>
- 4 Oxburgh, G. E., Myklebust, T., & Grant, T. (2010). The question of question types in
5 police interviews: a review of the literature from a psychological and linguistic
6 perspective. *International Journal of Speech, Language & the Law*, 17(1), 45-
7 66. 10.1558/ijssl.v17i1.45
- 8 Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., & R Core Team (2017). *nlme: Linear
9 and Nonlinear Mixed Effects Models. R package version 3.1-131*. Retrieved
10 from <https://CRAN.R-project.org/package=nlme>.
- 11 Powell, M., Roberts, K., & Guadagno, B. (2007). Particularisation of child abuse
12 offences: Common problems when questioning child witnesses. *Current issues
13 in criminal justice*, 19(1), 64-74.
14 <https://doi.org/10.1080/10345329.2007.12036415>
- 15 Price, H. L., & Connolly, D. A. (2013). Suggestibility effects persist after one year in
16 children who experienced a single or repeated event. *Journal of Applied
17 Research in Memory and Cognition*, 2(2), 89-94.
18 <https://doi.org/10.1016/j.jarmac.2013.03.001>
- 19 R Core Team. (2017). *R: A language and environment for statistical computing*. R
20 Foundation for Statistical Computing. Retrieved from [https://www.R-
21 project.org/](https://www.R-project.org/).

- 1 Roberts, A. (2003). Questions of who was there and who did what: The application of
2 code D when a suspect disputes participation but not presence. *Criminal Law*
3 *Review*, 709–716.
- 4 Roberts, W. T., & Higham, P. A. (2002). Selecting accurate statements from the
5 cognitive interview using confidence ratings. *Journal of Experimental*
6 *Psychology: Applied*, 8(1), 33-43. <http://dx.doi.org/10.1037/1076-898X.8.1.33>
- 7 Rubínová, E., Blank, H., Koppel, J., & Ost, J. (2020). Schema and deviation effects in
8 remembering repeated unfamiliar stories. *British Journal of Psychology*.
9 Advance online publication. <https://doi.org/10.1111/bjop.12449>
- 10 Schank, R. C., & Abelson, R. P. (1977). *Scripts, plans, goals, and understanding: An*
11 *inquiry into human knowledge structures*. Erlbaum.
- 12 Shepherd, E. & Griffiths, A. (2013). *Investigative Interviewing: The Conversation*
13 *Management Approach* (second edition). Oxford University Press.
- 14 Theunissen, T. P., Meyer, T., Memon, A., & Weinsheimer, C. C. (2017). Adult
15 eyewitness memory for single versus repeated traumatic events. *Applied*
16 *cognitive psychology*, 31(2), 164-174. <https://doi.org/10.1002/acp.3314>
- 17 Tullis, J. G., & Benjamin, A. S. (2015). Cueing others' memories. *Memory &*
18 *cognition*, 43(4), 634-646. <https://doi.org/10.3758/s13421-014-0478-y>
- 19 Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in
20 episodic memory. *Psychological Review*, 80(5), 352–
21 373. <https://doi.org/10.1037/h0020071>

- 1 Van der Vaart, W., & Glasner, T. (2007). Applying a timeline as a recall aid in a
2 telephone survey: A record check study. *Applied Cognitive Psychology*, *21*(2),
3 227–238. <https://doi.org/10.1002/acp.1338>
- 4 Weinsheimer, C. C., Coburn, P. I., Chong, K., MacLean, C. L., & Connolly, D. A.
5 (2017). Perceptions of credibility for a memory report of a single versus
6 repeated event. *Applied Cognitive Psychology*, *31*(4), 414-423.
7 <https://doi.org/10.1002/acp.3340>
- 8 Wheeler, R. L., & Gabbert, F. (2017). Using self-generated cues to facilitate recall: a
9 narrative review. *Frontiers in psychology*, *8*, 1830.
10 <https://doi.org/10.3389/fpsyg.2017.01830>
- 11 Willén, R. M., Granhag, P. A., Strömwall, L. A., & Fisher, R. P. (2015). Facilitating
12 particularization of repeated similar events with context-specific
13 cues. *Scandinavian Journal of Psychology*, *56*(1), 28-37.
14 <https://doi.org/10.1111/sjop.12180>
- 15 Woiwod, D. M., Fitzgerald, R. J., Sheahan, C. L., Price, H. L., & Connolly, D. A.
16 (2019). A meta-analysis of differences in children’s reports of single and
17 repeated events. *Law and human behavior*, *43*(1), 99-116.
18 <https://doi.org/10.1037/lhb0000312>
- 19 Yoshihama, M., Gillespie, B., Hammock, A. C., Belli, R. F., & Tolman, R. M. (2005).
20 Does the life history calendar method facilitate the recall of intimate partner
21 violence? Comparison of two methods of data collection. *Social Work Research*,
22 *29*(3), 151-163. <https://doi.org/10.1093/swr/29.3.151>

23

1 Table 1. Overview of superordinate recurring details across events and variations within
 2 each instance (in both deviation conditions).

Type of detail	Events			
	1	2	3	4
Target	The head of the National reform party	Swedish left-wing sociologist	Local activist	Investigative journalist
Attack Location	Spinnaker Tower	Portland building	Victoria park	Portsmouth & Southsea train station
	Base of the tower	Road in front of the building	North-west end of park	(Train arriving) from London
	Back of the café	Space under the stairs	Café lodge	Café inside the station
	Construction equipment	-	In bushes	Next to the sofa
	Seafront hotel	-	Premier hotel	Ibis hotel
	Two blocks away	The Hard	Closest bus stop	Square
	Within 900m	Within 900m	Within 900m	Within 800m
	Hotel lobby	-	Hotel lobby	Parking
Object/equipment	M48 Mortar rounds	3 kilos	Six packages	9 kilos
	-	C-4	Semtex	106 rounds
	Improvised Explosive Device	-	Improvised Explosive Device	Improvised Explosive Device
	Nokia 3210	Nokia 6210	Nokia 5210	Nokia 3310
	Backpack	Backpack	Briefcase	Sports rucksack
Time	Wednesday	Thursday	Saturday	Monday
	10 am	6 pm	12 o'clock	9 am
	10.30 am	6.10 am	Half an hour before	Two hours before

	6.30 am	-	Half an hour upon start (of the event)	10 minutes
	-	-	9 am	6 am
	Until the party starts	Until the reception starts	After the blast	After the blast
Operational roles				
Person planting bomb	Niko (male)	Nina (female)	Niko (male)	Nina (female)
Lookout	Taking pictures	Not taking pictures	Taking pictures	Not taking pictures
Passing the detonator	Nina hands detonator to Niko over discussion ⁷	Niko hands detonator to Nina ⁸	Nina hands detonator to Niko over discussion	Niko hands detonator to Nina

1
2
3
4
5
6
7
8
9
10
11
12

⁷ Nina asks Niko for directions and uses a map to conceal the detonator which she passes over to him

⁸ Niko hands the detonator to Nina without any interaction

1 Table 2. Protocol for follow-up open-ended in-depth questions to extend and clarify
2 information provided in the initial account.

1	Tell me more about (the part when/person/object/activity) ...
2	(You mentioned)...Tell me everything/ every detail about the part when ...
3	What else can you tell me about ...?
4	Explain in more detail what you mean about (this part where...)
5	Describe in more detail (this part when...)

3

4

5

6

7

8

9

10

11

12

13

14

15

16

17

18

19

1 Table 3. Mean and SDs accuracy rates of the initial reports and of the responses to the
2 follow-up questions across format conditions.

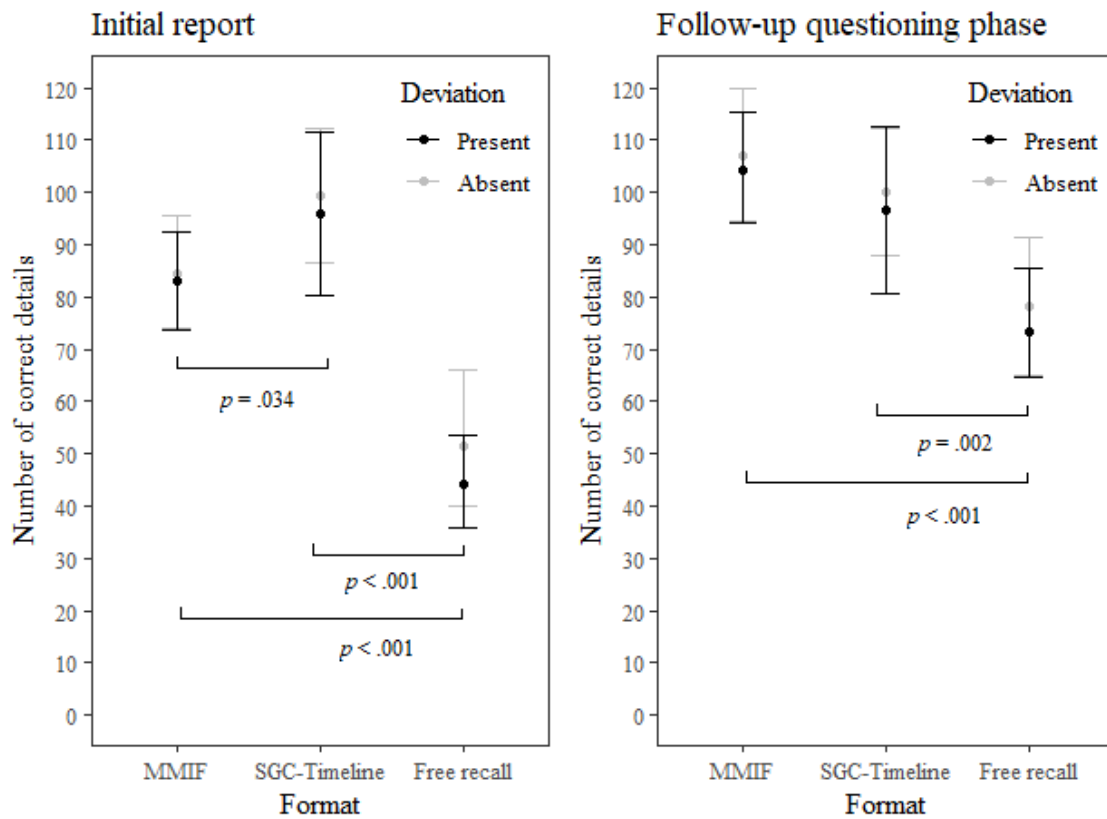
Format	Initial report		Follow-up questions	
	Mean Accuracy rate	<i>SD</i>	Mean Accuracy rate	<i>SD</i>
MMIF	0.87	0.06	0.78	0.10
SGC-Timeline	0.86	0.08	-	-
Free Recall	0.83	0.08	0.79	0.10

3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20

1 Table 4. Mean number and SDs of intrusions reported across format and deviation
 2 conditions.

	Deviation Absent			Deviation Present		
	MMIF	SGC-Timeline	Free Recall	MMIF	SGC-Timeline	Free Recall
Events	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>
First	.84 (1.28)	.96 (1.10)	.88 (1.20)	.84 (1.18)	.84 (.99)	.88 (1.01)
Second	.92 (1.08)	1.28 (1.17)	.92 (1.22)	1.00 (1.12)	.52 (.82)	.96 (.98)
Third	.96 (1.06)	.88 (.97)	.56 (.96)	1.52 (1.23)	.68 (.95)	.92 (1.26)
Fourth	1.08 (1.12)	.92 (.91)	.52 (.71)	.64 (.95)	.56 (.87)	.80 (.82)

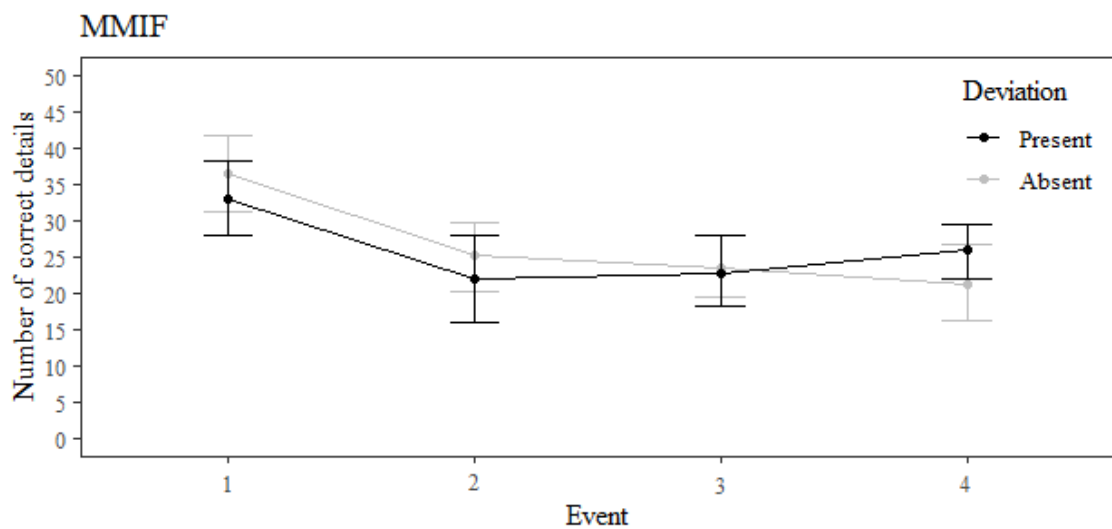
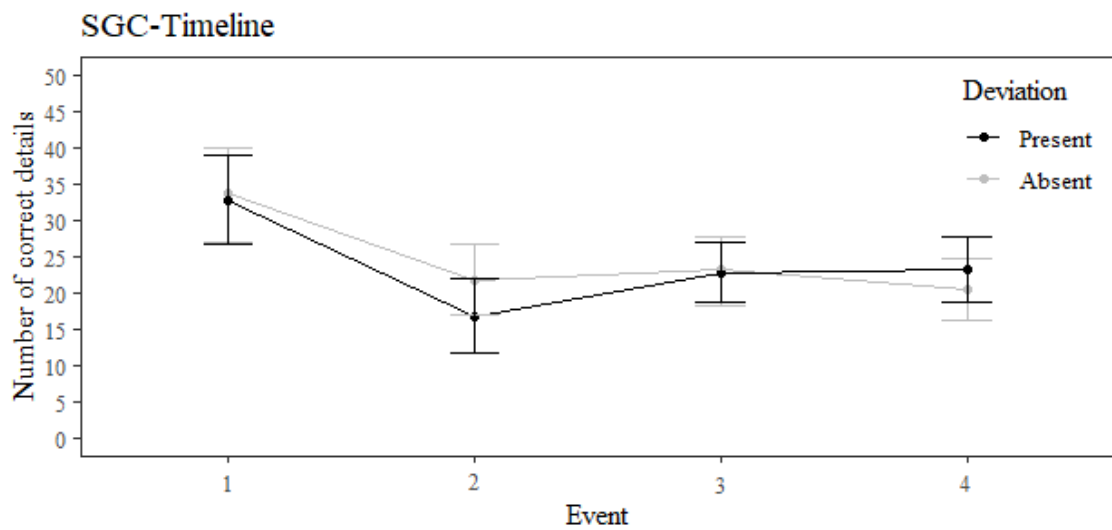
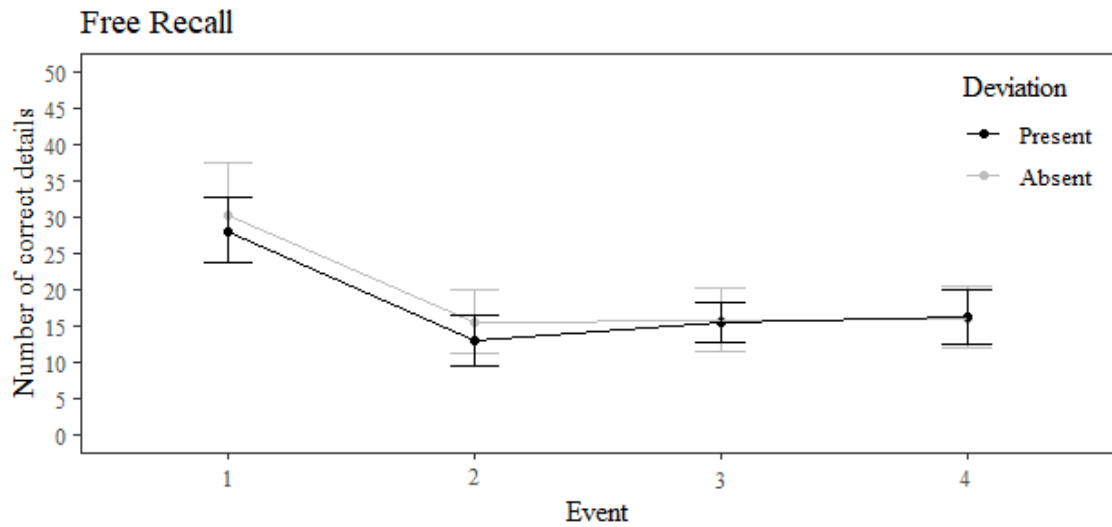
3
 4
 5
 6
 7
 8
 9
 10
 11



1

2 *Figure 1.* Number of correct details as a function of Reporting format within Present
 3 and Absent deviation conditions. Error bars represent 95% Confidence Intervals. Note:
 4 reporting in the SGC-Timeline condition remained stable (i.e., no follow-up questions
 5 were used).

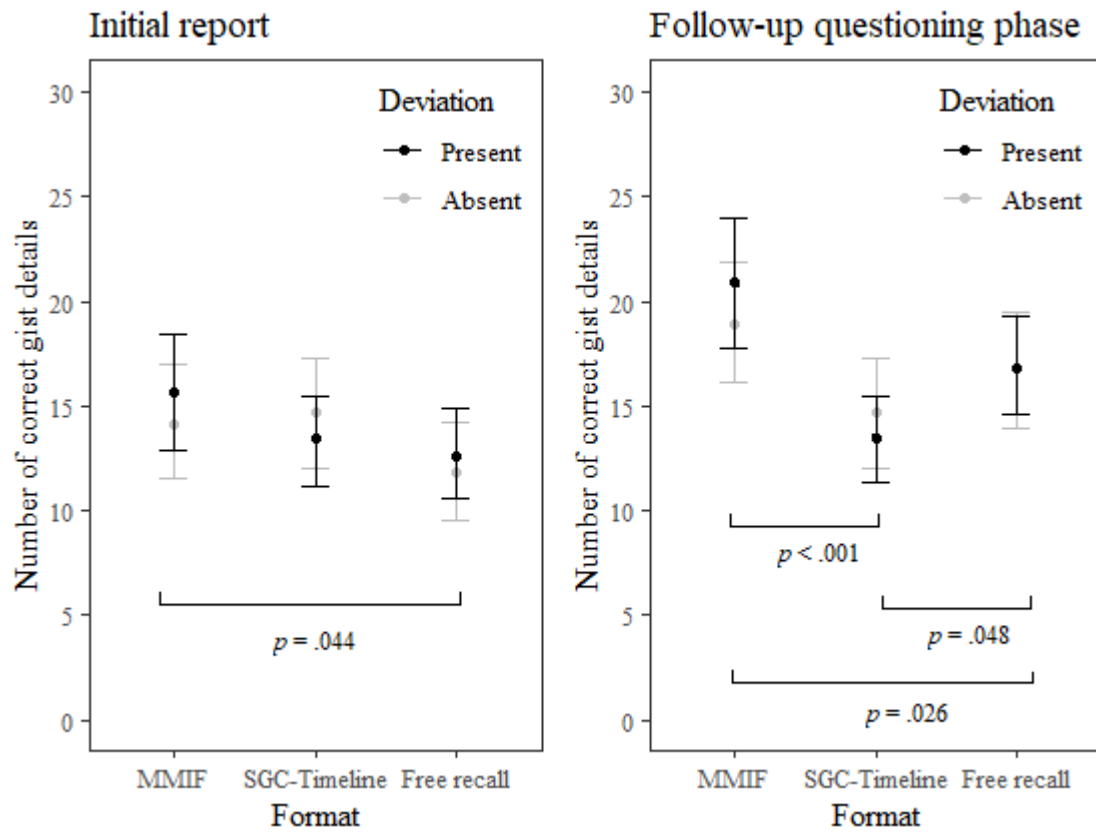
6



1

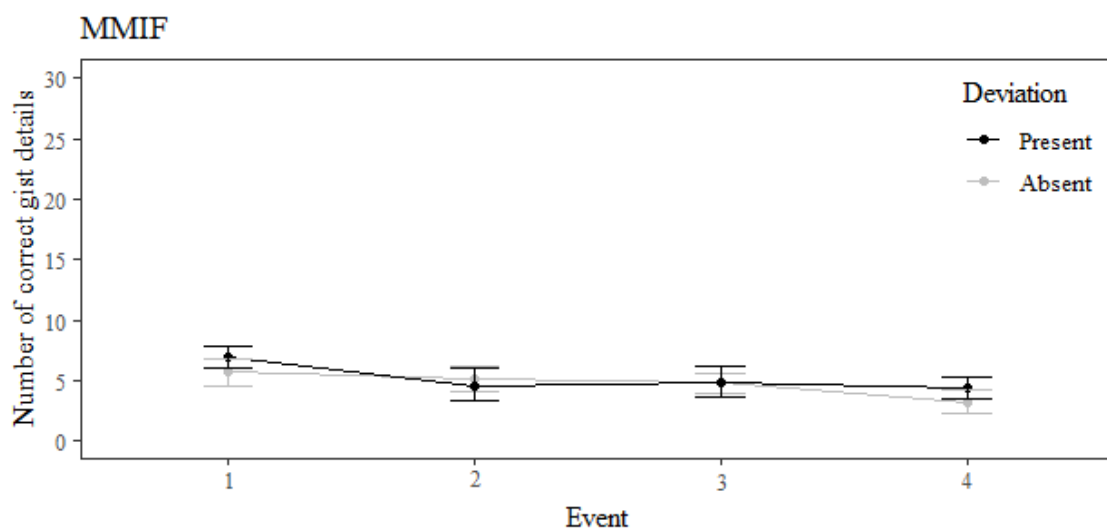
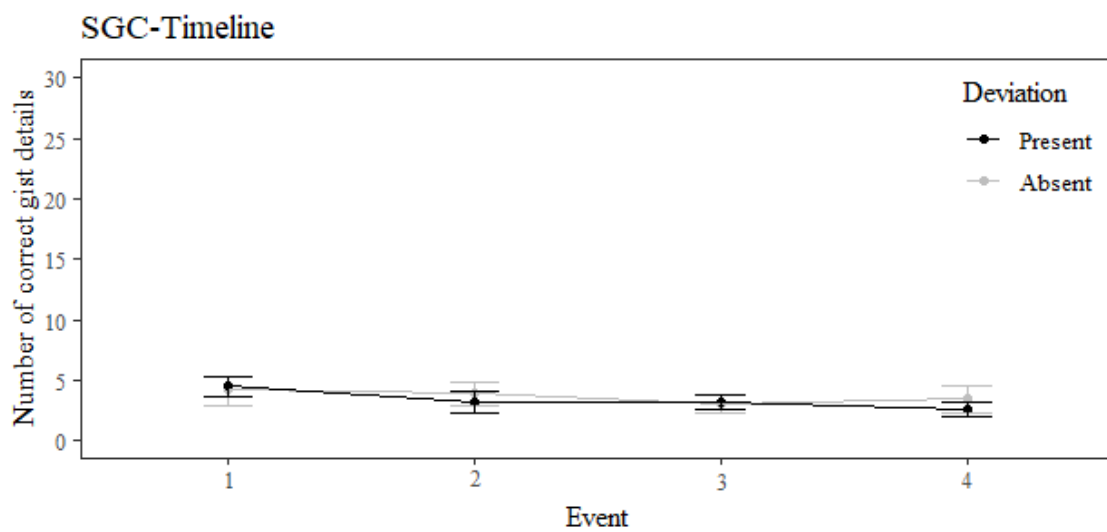
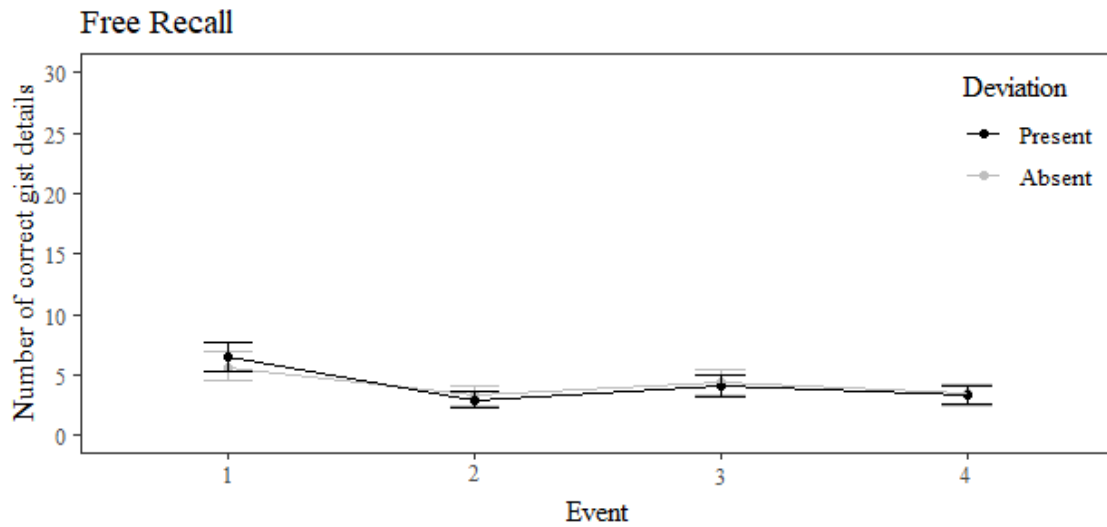
2 *Figure 2.* Number of correct details reported across events within Present and Absent
 3 deviation conditions for each format. Error bars represent 95% Confidence Intervals.

4



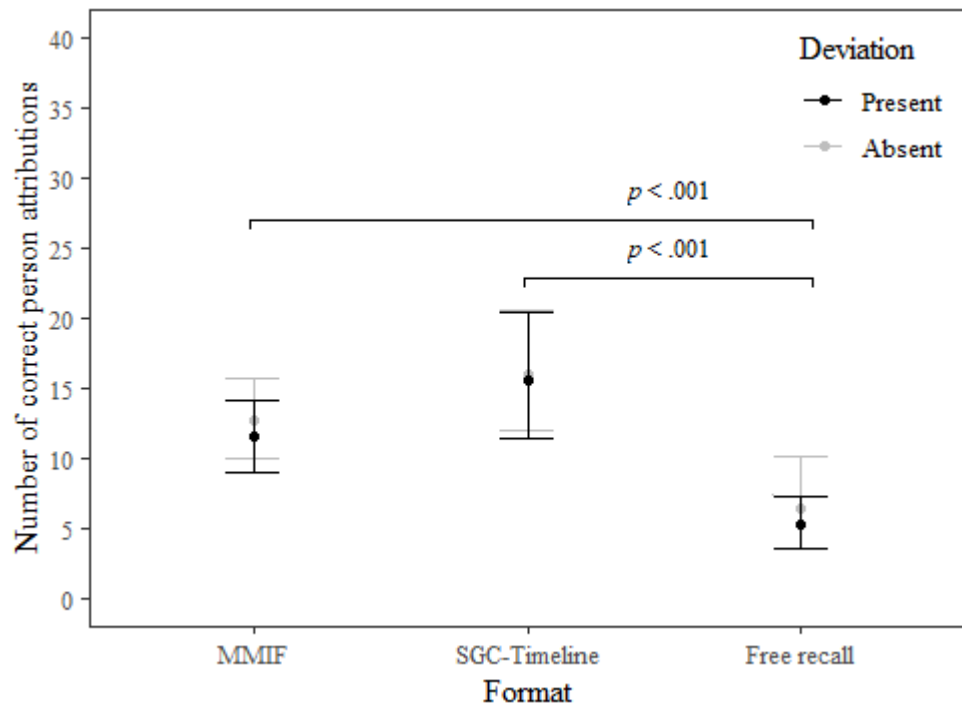
1

2 *Figure 3.* Number of correct gist details as a function of format within Present and
 3 Absent deviation conditions. Error bars represent 95% Confidence Intervals. Note:
 4 reporting in the SGC-Timeline condition remained stable (i.e., no follow-up questions
 5 were used).



1

2 *Figure 4.* Number of correct gist details reported across events within Present and
 3 Absent deviation conditions for each format. Error bars represent 95% Confidence
 4 Intervals.



1

2 *Figure 5.* Number of correct action and statement attributions to persons as a function of
 3 format within Present and Absent deviation conditions. Error bars represent 95%
 4 Confidence Intervals.

5

6

7