Notice: This is the author's version of a work that was accepted for publication in *Psychological Research.* A definitive version was subsequently published in *Psychological Research*. DOI: 10.1007/s00426-020-01310-7

Manuscript accepted for publication (22nd of February 2020) in *Psychological Research*

Noting a Difference: Change in Social Context Prompts Spontaneous Recall in 46-Month-Olds, But Not in 35-Month-Olds

Trine Sonne*[1]

Osman S. Kingo*[2]

Dorthe Berntsen*[3]

Peter Krøjgaard*[4]

*Center on Autobiographical Memory Research, Department of Psychology and Behavioral Sciences, Aarhus University, Denmark.

[1]= trines@psy.au.dk

[2]= osman@psy.au.dk

[3]= dorthe@psy.au.dk

[4]= peter@psy.au.dk

Corresponding Author:

Trine Sonne, Bartholins Allé 11, 1350-424, Center on Autobiographical Memory Research, Department of Psychology and Behavioral Sciences, Aarhus University, DK-8000 Aarhus C, Denmark

Phone: +45 87 16 53 61

E-mail: trines@psy.au.dk

Acknowledgements

We would like to thank our student assistants for their indispensable help getting this study done. More specifically, we would like to thank Amalie Juul Harregaard, Anne Rytter Christensen, Louise Jensen, and Sarah Jakobsen for testing the participants. Yibin Zhang for preparing the data for later coding, Caroline Beyer and Mette Sørensen for coding the data and finally Emilie Kristensen for preparing data for later analysis. Most importantly, we would like to thank all of the participating children as well as their parents for bringing them to our lab and for adhering to our comprehensive instructions in this design.

This study has been supported by the Danish National Research Foundation (DNRF89) and the VELUX FOUNDATION (VELUX10386).

Notice: This is the author's version of a work that was accepted for publication in *Psychological Research.* A definitive version was subsequently published in *Psychological Research*. DOI: 10.1007/s00426-020-01310-7

Abstract

An experimental paradigm has shown that it is possible to activate spontaneous memories in children by having them re-visit the setting in which they were introduced to a memorable event. Nevertheless, the most important cues for spontaneous recall remain undetermined. In response, we investigated the importance of the experimenter by introducing 35-month-olds (n = 62) and 46-month-olds (n = 62) to the same or a new person after one week. We expected that altering the experimenter would result in fewer recollections through reducing the overlap of cues between encoding and testing. In contrast, the manipulation affected the two age groups differently: No effect of Condition was seen in the 35-month-olds, whereas the 46-month-olds performed better, when the experimenter had changed, suggesting a sensitivity to change and an ability to update their knowledge of the event. We replicated previous findings demonstrating that both age groups exhibited spontaneous recollections.

Keywords: spontaneous recall; experimenter; cue; memory; involuntary recall

Notice: This is the author's version of a work that was accepted for publication in *Psychological Research.* A definitive version was subsequently published in *Psychological Research*. DOI: 10.1007/s00426-020-01310-7

Noting a Difference: Change in Social Context Prompts Spontaneous Recall in 46-Month-Olds, But Not in 35-Month-Olds

1. Introduction

Asking a child to tell you about what happened on a given day in daycare may seem like a basic request, but it actually constitutes a cognitively demanding task for young children, potentially resulting in a lack of detailed information. Indeed, responding to direct questions requires deliberate and strategic recall, which is in itself a challenge for young children (Krøjgaard, Kingo, Jensen, & Berntsen, 2017), probably because the prefrontal cortex is involved in strategic retrieval, but frontal lobes only mature late in ontogenesis (e.g., Johnson, 2005).

Nevertheless, strategic retrieval is not the only path to memory. At times, memories come to us uninvited, suddenly and almost 'out of the blue'. Research focusing on adults has dubbed this form of memory retrieval *involuntary recall* (e.g., Berntsen, 2009). It is predominantly associative and hence less cognitively demanding, and has been suggested to be a basic mode of remembering that may developmentally precede voluntary recall (e.g., Berntsen, 2009, 2012).

When carefully instructed, adults can distinguish whether they thought of a memory involuntarily or voluntarily. In contrast, young children are unlikely and probably unable to make such metacognitive distinctions. Therefore, instead of employing a first-person perspective, in which the participant makes his or her own assessment of how a memory was retrieved, developmental researchers have moved to a third-person perspective (allowing the researcher to make the distinction). The memories are defined as *spontaneous*, characterized as being verbally produced, non-prompted, and typically environmentally cued (Krøjgaard, Kingo, Dahl, & Berntsen, 2014; Krøjgaard et al., 2017). In a recent series of experiments, Krøjgaard and colleagues have provided evidence that children just below the age of three can have spontaneous recollections even though they struggle to strategically recall the event (Krøjgaard et al., 2014; Krøjgaard et al., 2017; Sonne, Kingo, Berntsen, & Krøjgaard, 2019). While the paradigm developed by Krøjgaard et al. (2017) has been used to successfully induce spontaneous recall in young children, the criterion that a spontaneous memory should be verbally produced results in an age threshold when testing children using this paradigm. Other paradigms for testing infant memory (e.g., the Visual Paired-Comparison test or the imitation paradigm) have in some sense also been used to document spontaneous (that is, 'unprompted') memory even in infancy (for an overview see Hayne, 2004). Nevertheless, Krøjgaard et al.'s (2017) paradigm was developed as a test of explicit memory to contrast the two types of retrieval. Further, this paradigm places considerable emphasis on the criterion that spontaneous utterances should be 'unprompted' and thus not the result of statements such as "Can you show me how to play with these objects?", as widely used in the imitation paradigm (e.g., Sonne, Kingo, & Krøjgaard, 2016).

The present study is based on Krøjgaard et al.'s (2017) design, in which 35- and 46-month-old children were presented with one of two engaging and interactive events: the first concerning mechanical singing teddies; the other involving two games for which the children would win prizes no matter how they performed. After one week, the children returned to the laboratory, and this time they and their parents were left alone in a room while everything the children might say was recorded. This two-minute waiting period was compared to a baseline measure collected during the first visit prior to exposing the children to one of the two engaging events. The rationale was that returning to the same setting (including the room and experimenter) would facilitate substantial cuing, which was likely to result in spontaneous recall of the event (Krøjgaard et al., 2017). The results revealed that children in both age groups spontaneously remembered the previously experienced event, and that overall the 35-month-olds performed just as well as the 46-month-olds.

Notice: This is the author's version of a work that was accepted for publication in *Psychological Research.* A definitive version was subsequently published in *Psychological Research*. DOI: 10.1007/s00426-020-01310-7

From both the literature on adults' involuntary memories (e.g., Berntsen, 2009; Berntsen, Staugaard, & Sørensen, 2013) and diary studies with children (e.g., Reese, 1999), we know that spontaneous recall is typically triggered by cues in the environment. For instance, when passing by a bakery, the smell of freshly baked bread might suddenly make you think of a Sunday afternoon when as a child you tasted your mother's buns hot out of the oven. The encoding specificity principle (e.g., Tulving & Thomson, 1973) suggests that for cues to be effective, there must be an overlap with the circumstances during encoding and the time of retrieval, in this case the aroma of freshly baked bread. However, overlap between encoding and testing is not enough on its own: the overlap should be sufficiently *distinct* to work as a cue for a specific memory (see e.g., Berntsen, 2009; Hunt & Smith, 1996; Nairne, 2002).

In the study by Krøjgaard et al. (2017), spontaneous recall was induced by testing the children in the exact same setting as where encoding occurred, involving several potentially distinct cues: the same room, same furnishing and the same experimenter. However, given that the contextual overlap was complete, we cannot identify which specific cues triggered children's spontaneous recall: was it the room, the furnishing, the boxes, the experimenter, or a combination of these cues? Clarification of the cues that trigger spontaneous recall is critical for our understanding of the associative mechanisms involved in this type of retrieval. Furthermore, it would provide knowledge regarding context dependency in spontaneous recall. Note that the majority of previous studies documenting and examining spontaneous recall in children have been diary studies (Nelson & Ross, 1980; Reese, 1999; Todd & Perlmutter, 1980), in which meticulous experimental examination of candidate cues is beyond reach. In contrast, the recently developed experimental design outlined above allows for systematic examination of the potential impacts of different kinds of cues.

A recent study from our lab using the same overall design explored this possibility, specifically concerning the significance of the location for retrieval in 35-month-old children. The results revealed that although changing the location (to a novel and somewhat differently furnished room) of the second visit led to numerically reduced spontaneous recall relative to a control group that returned to the same room, these differences were not significant, and the children in the experimental group still reliably recalled the event spontaneously (Sonne et al., 2019). While changes in the overall context did not appear to affect the children's spontaneous recollections, other cues remained constant between encoding and testing, potentially working as memory cues and hence triggering spontaneous recall, for example, the experimenter, the boxes or the furnishing of the room.

Although the experimenter obviously is an important aspect in most laboratory studies, his or her importance as a cue has rarely been examined (for a similar argument see Goldenberg & Sandhofer, 2013), especially beyond studies of infant memory. Studies employing the deferred imitation paradigm have for instance demonstrated how changing the experimenter at test can have detrimental effects on infants' abilities to imitate action sequences (e.g., Learmonth, Lamberth, & Rovee-Collier, 2005). Learmonth et al. (2005) investigated infants' (6-, 9-, 12-, 15- and 18-month-olds) ability to participate in a deferred imitation task, noting that in instances in which the experimenter was new at test, infants were unable to imitate the actions. Others have found that infants (14- to 18-month-olds) *can* imitate even when tested by a new experimenter, albeit they did imitate fewer actions on average compared to infants with no change in context (Hanna & Meltzoff, 1993).

Studies with this focus in relation to children are almost non-existent. One investigation examined whether changing the experimenter would affect 3-, 4- and 5- year-olds' categorization (Goldenberg & Sandhofer, 2013). In this experiment, the children were first introduced to labels for different object categories and subsequently tested in a forced choice generalization task. Interestingly, Goldenberg and Sandhofer (2013) found that although the 3-year-olds performed similarly regardless of whether the experimenter was the same or a new person at test, 4- and 5-year-olds performed worse when the experimenter was new. The finding that the youngest age group was not negatively affected by the change was explained by poor source monitoring skills at this age. This literature makes it evident that very little is known about the role of the experimenter as a potential cue, and no previous study has examined the role of the experimenter as a cue for spontaneously arising memories in young children.

1.1 The present study

In the present study, we set out to investigate the importance of the experimenter as a cue for subsequent recall, with the broader aim of specifying the cues that may trigger spontaneous retrieval in young children. We used the procedure developed by Krøjgaard et al. (2017), but added the manipulation that half of the children would meet a new experimenter in the second visit (Change of Experimenter Condition), whereas the other half would return to the same experimenter at test (Same Experimenter Condition). We decided to test the same two age groups as in Krøjgaard et al. (2017) – 35-month-olds and 46-month-olds – in order to relate our results to the reference study.

First, based on previous studies using the same design, we expected that the children (in both age groups) would have spontaneous recollections. Second, based on the encoding specificity principle and the importance of distinct cues for triggering spontaneous recall, we expected that the absence of the same experimenter in the Change of Experimenter Condition at test would make it less likely for the children to spontaneously recall the event, such that this group of children would have fewer recollections relative to the group of children returning to the same experimenter. We had no firm hypothesis regarding age differences in relation to this manipulation,

Notice: This is the author's version of a work that was accepted for publication in *Psychological Research.* A definitive version was subsequently published in *Psychological Research*. DOI: 10.1007/s00426-020-01310-7

because no previous studies have altered the experimenter as regards verbal spontaneous recollections. Nevertheless, based on the results of Goldenberg and Sandhofer (2013), age differences might be expected.

2. Method

2.1 Participants

A total of 124 children participated in the study. Half of these children were 35-month-olds (31 girls, M_{age} = 35.19 months, SD = .50, range 34.4-36.7) and half were 46-month-olds (35 girls, M_{age} = 46.47 months, SD = .85, range 44.1-48.1). All of the participating children were healthy and full-term recruited from birth registries from the National Board of Health in BLINDED. The children were predominantly Scandinavian Caucasian and from families with medium to high socio-economic status. At the second visit, the children received a small present for participating. No further compensation was offered for participation.

To test the importance of the experimenter as a cue, the children were randomly assigned to one of two conditions: *the Change of Experimenter Condition* (n = 62; thirty-one 35-month-olds, and thirty-one 46-month-olds), in which the children would meet a different experimenter when returning for the Test session (T_2); and *the Same Experimenter Condition* (n= 62; thirty-one 35-month-olds, and thirty-one 46-month-olds), whereby the children would meet the same experimenter at both the Encoding session (T_1) and at the Test session (T_2). Furthermore, within each condition, half of the children were presented with a Teddy event, and the other half with a Game event (see description below). An additional 13 children were tested but excluded due to any of the following reasons: fussiness = 5; experimental error = 1; parental interference/parents not adhering to our guidelines = 4; speaking a foreign language at test = 1; technical errors = 2.

The study was approved by the local ethics committee at the Department of Psychology at BLINDED University. Written and informed consent was obtained.

2.2 Materials

Testing took place in a sparsely furnished 16 m^2 room. Two mechanical teddies were used for the *Teddy event*: Elly, an elephant that could sing and wiggle its ears; and Alfred, a dog that was able to sing, clap and wiggle its ears simultaneously. The teddies were commercially available and imported from abroad for the purpose of the study, in order to avoid the risk of the children already being familiar with them. For the *Game event*, two different games were used: a home-made throwing game in which children had to throw three balls into three buckets; and a bowling game in which the children had to make the pins turn over.

To ensure that the children could not see the props used in the events during the two-minute waiting period, all of the props were constantly locked away in boxes, except for when being used for demonstration. For the Teddy event, the props were hidden in a red metal box; for the Game event, the props were locked away in a grey plastic box normally used for cushions for outdoor use.

2.3 Design and procedure

All of the children visited the lab twice, with a one-week retention interval (*M* = 6.98 days, *SD* = .49, range = 6-9) between the two visits, following the design of Krøjgaard et al. (2017). During the second visit, half of the children were met by a new experimenter. In total, five different female experimenters (all in their twenties or early thirties) tested the children. They varied in several aspects of their physical appearance: one was pregnant, one wore glasses, and they had different hair colors (varying from blonde to dark blonde to red). Thus, they were all easily distinguishable. All experimenters had been carefully instructed to follow the same manualized protocol stating exactly what to say and do during the experiment. They had been trained using video material from previous studies employing the same paradigm, and they were taught to use the same language and to make sure that each event demonstration lasted approximately six minutes. The parents had been carefully instructed by mail and phone before the visit to ensure that they would not inform their children that the study was about memory, and that they should therefore not initiate conversation with them during the two waiting periods. Furthermore, they were asked not to talk about the event at home during the one-week retention period.

2.3.1 Encoding (T_1)

In the first visit, the child and parent were collected from the waiting room. The parent was reminded via a written note not to initiate conversation or to follow up on their child's comments. They were then brought to the lab, and as soon as written consent was obtained, the experimenter excused herself by saying that she had to quickly take care of something and would return soon. The child and the parent then waited alone for the two-minute baseline period and everything was recorded for later coding. Following the baseline period, the experimenter re-entered the room and presented the child with one of the two events. Both events had a duration of about six minutes. For the *Teddy event*, the experimenter would open the red metal box and demonstrate to the child how the teddies could sing, clap and wiggle their ears. The child was also encouraged to participate by pressing the buttons that made them sing as well as to name the animals. For the *Game event*, the experimenter opened the grey box and presented the child first with the throwing game and then with the bowling game. Again, the children in this event were asked to name the props and to participate in the games. Furthermore, the children received a 'gold' medal for their participation.

Notice: This is the author's version of a work that was accepted for publication in *Psychological Research.* A definitive version was subsequently published in *Psychological Research*. DOI: 10.1007/s00426-020-01310-7

After the visit, the parents received an e-mail with a link to an electronic version of the BLINDED MacArthur-Bates Communicative Development Inventory: Words and Sentences. The parents were asked to complete it before returning for the second visit.

2.3.2 Test of spontaneous recall (T_2)

When the children and their parent(s) returned for the second visit, they were again met in the waiting room, brought to the lab, and then left alone for a two-minute waiting period. Half of the children returned to a new experimenter, and crucially, all of the children now knew the contents of one (but only one) of the boxes. Aside from this, everything else was identical to the baseline period. In the second visit before the test, the children were thus only together with the (new or familiar) experimenter during the couple of minutes required for them to walk from the waiting room to the lab. After the two-minute waiting period, the experimenter returned to ask each child questions regarding a different aspect of the study in order to assess his or her strategic recall of the event (to be reported elsewhere).

At the end of the session, the children were asked a question aimed at investigating whether those in the Change of Experimenter Condition noticed the change. Specifically, all of the children were asked:

Q: So at the previous visit you were presented with something inside that box [experimenter points at the target box]. Did *I* show it to you?

This question was added to see whether the children indeed noticed the change. Adding this question represented an improvement of the study by Sonne et al. (2019) in which the children returned to a different room at test, but where it was unclear whether they actually noticed this change, because they were never asked about it.

Notice: This is the author's version of a work that was accepted for publication in *Psychological Research.* A definitive version was subsequently published in *Psychological Research*. DOI: 10.1007/s00426-020-01310-7

2.4 Coding

We used the same coding strategy as described by Krøjgaard et al. (2017) and Sonne et al. (2019). Therefore, to code any spontaneous utterance, two videos from the two-minute waiting time periods were created for each participant: one from the first visit and one from the second visit. Importantly, the coders were unaware of which event the children had been presented with, whether the video was from the first or the second visit, and the hypotheses of the study. Coding was undertaken by a primary coder, as well as a secondary coder to re-code a representative (i.e., from both conditions and time points) 20 percent of the data. Inter-rater agreement was high: 98.7 percent (range: 90.5-100%).

Spontaneous utterances were coded by two means: a Word List and a Coding Scheme. These will be briefly described.

2.4.1 Word List

Video clips from both the first and second visits were coded by use of a predefined Word List (see Table 1). The Word List was based on words that the children were introduced to during the demonstration (based on Krøjgaard et al., 2017), but in the present study also included words that we expected that the children might say as a result of changing the experimenter (e.g., 'new girl'). The Word List consisted of event-specific words (concerning the Teddy or the Game event) as well as unspecific words referring to the previous visit but without explicit reference to either event (e.g., "She did that at the previous visit too"). Each verbal utterance was coded in relation to each of these categories. Sum scores were calculated for each child.

2.4.2 Coding Scheme

The utterances were then analyzed through a Coding Scheme in order to cast a wider net and capture any sign of spontaneous recall that the Word List analysis may have missed. Every two-minute video clip was divided into 12 time slots (10s each), and for each of these the coders had to respond to six specific questions pertaining to different dimensions (inspired by Levine, Svoboda, Hay, Winocur, & Moscovitch, 2002).[1] The questions were as follows:

1. Does the child by means of *language* refer to knowledge that originates from a previous visit?
2. Does the child by means of *gestures* refer to knowledge that originates from a previous visit?
3. Does the child's verbal and/or non-verbal behavior indicate that he or she mentally *re-lives* parts of a previous visit?
4. Does the child refer to specific *action details* from a previous visit?
5. Does the child by means of specific *spatial details* refer to knowledge that originates from a previous visit?
6. Does the child by means of specific *social details* refer to knowledge that originates from a previous visit?

Within each 10s slot, a question could give a single score or no score. A sum score was created based on the results of the twelve 10s slots (range: 0-12) for each dimension. For every child we thus attained six sum scores, one for each dimension.

For instance, one child said: "We will try that game [pointing to the grey box] when the lady comes back!" This statement resulted in a score on the dimensions 1, 2, 3, 4, and 6 (but not 5). Another child just stated: "Dog, elephant!" This statement resulted in a score on the dimensions 1 and 3 (but not 2, 4, 5, and 6).

2.4.3 Question concerning change of experimenter

[1] In Krøjgaard et al. (2017), seven dimensions were used, but in the present study we decided to omit the 'object' dimension because we encountered problems in identifying unambiguous criteria by which the object dimension could be coded during coder training.

Notice: This is the author's version of a work that was accepted for publication in *Psychological Research.* A definitive version was subsequently published in *Psychological Research*. DOI: 10.1007/s00426-020-01310-7

For the final question concerning the change in experimenter, the experimenter during testing simply noted whether the answer was correct or not. Given that coding was straightforward, re-coding was not deemed necessary.

3. Results

3.1 Preliminary analyses: MacArthur-Bates Communicative Development Inventory

As a preliminary analysis, we first wanted to check for potential differences in relation to the vocabulary scores. Running an ANOVA with Age Group (35 vs. 46 months) as a between-subjects factor and children's productive vocabulary scores as a dependent variable, as expected we found a main effect of Age Group: $F(1, 121) = 11.156$, $p = .001$, $\eta_p^2 = .084$. Unsurprisingly, the 46-month-olds (M = 629.80, SD = 102.71) possessed a larger vocabulary than their younger peers (M = 568.87, SD = 99.62).[2] In addition, when looking at the different age groups separately, by running an ANOVA with Condition (Same Experimenter vs. Change of Experimenter) and Event (Teddy vs. Game) as between-subjects factors and children's productive vocabulary scores as the dependent variable, no differences were found between the two conditions or events. For the 35-month-olds, all ps >*.055*; and for the 46-month-olds, all ps >.119.

3.2 Word List and Coding Scheme

Our primary interest was to investigate two specific questions: 1) Did the children overall have spontaneous recollections? That is, at test (T_2) did they produce spontaneous utterances concerning the event that they had experienced at T_1? 2) Did the change in experimenter affect the results?

To this end, we were interested in whether the spontaneous utterances produced by the children were congruent with the event (Teddy or Game) they experienced during the first visit. A

[2] Given that we failed to obtain CDI data from one participant, these analyses were based on n = 123.

congruent spontaneous utterance was operationalized as referring specifically to the event (Teddy or Game) with which the child was presented at T_1, or by an *un*specific utterance, indicating that he or she had been there before, without referring to a specific event. An incongruent utterance, on the other hand, refers to the event that the child did not witness before. Given that this was our overall focus, and because the two events had been counter-balanced across the children and the experimental manipulation, the two events (Teddy or Game) were collapsed in the following analyses (following Krøjgaard et al., 2017; Sonne et al., 2019).

Overall, the children presented no scores pertaining to either congruent or incongruent utterances at T_1, and no scores in relation to *in*congruent utterances from the Test at T_2. Thus, as expected, none of the children discussed the events at the first visit before being presented with the contents of one of the boxes, and none of the children at the second session talked about the event they had never witnessed before. Therefore, in the following analyses our focus was on the congruent utterances that the children made at Test (T_2). Based on this, one-sample *t*-tests were tested against the value '0' in order to examine whether T_2 utterances differed from the baseline. Table 2 presents the descriptive statistics as well as *t*-tests from both events combined.

The results displayed in Table 2 show that the children reliably produced spontaneous recollections of the previously experienced event. In relation to the Word List, the children clearly had spontaneous recollections of the experienced event, and for the Coding Scheme dimensions, the same seemed to be the case; only the social dimension in the Coding Scheme was not significant, akin to the findings of Sonne et al. (2019), and in one instance, the spatial dimension was not significant. These results were systematic and generally replicated previous findings (Krøjgaard et al., 2017).

As in previous studies, the dimensions from the Coding Scheme could potentially be treated as a scale. Therefore, as in Krøjgaard et al. (2017), we ran an item analysis that yielded acceptable internal consistency in terms of Cronbach's α = .806. This was a little lower than in the previous study (α = .917), albeit still justifying calculating sum scores. However, we opted for a sum score comprising five of the six dimensions, given that removing the Social dimension increased the Cronbach's α (= .838). We thus calculated a sum score for all of the five dimensions and analyzed whether the results would replicate when running the same analyses with this sum score, which they did (see Table 2). To simplify the following analyses, we only used the sum score as a proxy for the dimensional measures.

3.3 Experimental manipulation and developmental changes

Having demonstrated that the children experienced spontaneous memories, we were interested in examining whether our manipulation regarding the experimenter affected their spontaneous recall. To this end, we ran two ANOVAs with Condition (Same Experimenter vs. Change of Experimenter) and Age Group (35 vs. 46 months) as between-subjects factors: one ANOVA with the results from the Word List as the dependent variable, and another ANOVA with the sum score from the Coding Scheme as the dependent variable. The first ANOVA based on the Word List measures resulted in no main effects (Condition: $F[1, 120] = 0.403$, $p = 0.527$, $\eta_p^2 = .003$; Age Group: $F[1, 120] = 1.031$, $p = 0.312$, $\eta_p^2 = .009$). Moreover, we found no significant interaction between Condition and Age Group, $F(1, 120) = 2.321$, $p = 0.130$, $\eta_p^2 = .019$ (see Table 2 for *M*s and *SD*s).

Running the same ANOVA, but this time inserting the sum score from the Coding Scheme as the dependent variable, the analysis again resulted in no main effects (Condition: $F[1, 120] = 0.00$, $p = 1.00$, $\eta_p^2 = .00$; Age Group: $F[1, 120] = 1.901$, $p = 0.171$, $\eta_p^2 = .016$). However, this

Notice: This is the author's version of a work that was accepted for publication in *Psychological Research.* A definitive version was subsequently published in *Psychological Research*. DOI: 10.1007/s00426-020-01310-7

time the interaction between Condition and Age Group was significant: $F(1, 120) = 7.21$, $p = .008$, $\eta_p^2 = .057$ (see Figure 1).

In order to explore the significant interaction between Age Group and Condition on the Coding Scheme sum score, we ran two ANOVAs (one for the Change of Experimenter Condition and another for the Same Experimenter Condition) with Age Group (35 vs. 46 months) as between-subjects factors and the results of the sum score from the Coding Scheme as the dependent variable. Age Group had no effect in the ANOVA for the Same Experimenter Condition ($F < 1$), which is in accordance with the previous study using this design. However, we did obtain a simple effect of Age Group in the Change of Experimenter Condition $F(1, 60) = 9.173$, $p = .004$, $\eta_p^2 = .133$ – as the 46-month-olds yielded reliably higher sum scores ($M_{46\_sum\_score} = 3.00$, $SD = 2.58$) than their younger peers ($M_{35\_sum\_score} = 1.19$, $SD = 2.09$).

When broken down by Age Group, and running two ANOVAs (one for each age group) with Condition (Same Exp. vs. Change of Exp.) as between-subjects factors and the results of the sum score from the Coding Scheme as the dependent variable, we obtained a simple effect of Condition for the oldest age group ($F[1, 60] = 4.029$, $p = .049$, $\eta_p^2 = .063$, see Table 2 for *M*s and *SD*s) and no significant difference in the youngest age group ($F[1, 60] = 3.262$, $p = .076$, $\eta_p^2 = .052$). Inspecting the means, we see that whereas changing the experimenter resulted in numerically (although not significantly) lower sum scores on the Coding Scheme measures for the youngest age group relative to the Same Experimenter Condition, the change of experimenter resulted in *higher* sum scores for the oldest age group in the Change of Experimenter Condition relative to the Same Experimenter Condition (see Figure 1).

Notice: This is the author's version of a work that was accepted for publication in *Psychological Research.* A definitive version was subsequently published in *Psychological Research*. DOI: 10.1007/s00426-020-01310-7

3.4 Additional analyses

One may speculate whether the (surprising) greater amount of spontaneous recall among the 46-month-olds in the Change of Experimenter Condition (Relative to the Same Experimenter Condition) as measured by the sum score from the Coding Scheme was driven by children who explicitly commented on this change, thereby resulting in more incidents of *unspecific* spontaneous utterances (e.g., "The lady is new", or "Why is it not the same girl?"). Among the thirty-one 46-month-olds participating in the Change of Experimenter Condition, seven children explicitly commented on the change of experimenter during the Test of spontaneous recall. In contrast, none of the 35-month-olds explicitly mentioned this change in the two-minute test. To assess whether these seven children might have driven the obtained effect of Condition among the 46-month-olds, we conducted an ANOVA with Mention Change (mentioned change of experimenter vs. did not mention change of experimenter) as between-subjects factor and the sum scores from the Coding Scheme as the dependent variable. Although this analysis did not result in a significant effect of the Mention Change factor, the means show that the 46-month-olds who commented on the change of experimenter had numerically higher sum scores ($M_{\text{mention\_change}}$ = 4.57, *SD* = 2.44) than their equally aged peers who did not mention the change ($M_{\text{no\_mention\_of\_change}}$ = 2.54, *SD* = 2.48), $F(1, 29) = 3.64$, $p = .066$, $\eta_p^2 = .112$.

To elaborate further, we also re-ran the ANOVA with Condition (Same Exp. vs. Change of Exp.) as between-subjects factor and with the sum scores from the Coding Scheme as the dependent variable, but this time *excluding* the seven children who mentioned the change of experimenter from the sample of 46 month-olds. This analysis revealed no effect of Condition, $F(1, 53) = 1.431$, $p = .237$, $\eta_p^2 = .026$ (Same Exp. Condition: $M_{\text{46\_sum\_score}}$ = 1.81, *SD* = 2.07; Change of Exp. Condition (reduced sample: n = 24): $M_{\text{46\_sum\_score}}$ = 2.54, *SD* = 2.48). Taken together, the results from these additional analyses converge by demonstrating that the responses of the seven 46-

month-olds who explicitly commented on the change in experimenter contributed substantially to the overall results.

To further pursue this effect, it was important to consider the fact that the sum scores from the Coding Scheme were based on both condition-specific (e.g., "There are teddies inside") and *un*specific scores (e.g., "I think she will open up the boxes again"), following previous studies using the same design (Krøjgaard et al., 2017, Sonne et al., 2019). The fact that the experimenter differed at test may be viewed as an *un*specific cue related to the overall scenario, not being specifically related to the Teddy or Game event as such. Thus, it is possible that a disproportionally large number of unspecific scores were responsible for the increase in spontaneous utterances among the 46-month-olds in the Change of Experimenter Condition. To test this possibility, we re-ran the ANOVA for the 46-month-olds with Condition (Same Exp. vs. Change of Exp.) as the between-subjects factor, but this time only with the sum scores from the *condition-specific* results from the Coding Scheme (that is, with the *un*specific scores left out) as the dependent variable. This analysis revealed that although the children in the Change of Experimenter Condition still obtained numerically higher sum scores ($M_{46\_Condition\_Specific}$ = 2.03, SD = 2.24) than their equally aged peers in the Same Experimenter Condition ($M_{46\_Condition\_Specific}$ = 1.29, SD = 2.00), the difference was no longer significant: $F(1, 60) = 1.887$, $p = .175$, $\eta_p^2 = .030$. Thus, the elevated level of sum scores in the Change of Experimenter Condition among the 46-month-olds was at least in part driven by an increase in unspecific spontaneous utterances, spurred by the change in experimenter, which probably captured their interest and attention.

3.5 Did they notice the change?

Finally, at test the experimenter asked the children whether she was the person who had shown them what was inside the box during the previous visit. The purpose was to examine whether the children in the Change of Experimenter Condition noticed the change. Among the 35-month-old children in the Change of Experimenter condition, only six out of 28 valid replies (excluding cases where children said 'I don't know' or did not answer the question) were correct, significantly below the chance level (binomial test, p = .004). For the 46-month-olds, 20 out of 29 valid responses (again excluding the cases where children said 'I don't know' or did not answer the question) were correct, although they did not perform above chance (p = .061).

For the 35-month-olds in the Same Experimenter Condition, 20 of 22 valid responses (again excluding cases where children said 'I don't know' or did not answer the question) were correct, whereas for the 46-month-old children, 28 of 30 valid responses (again excluding cases where children said 'I don't know' or did not answer the question) were correct. In this condition, the children performed above chance (in both cases p <.001).

4. Discussion

Studies have shown that spontaneous memories of previously experienced events can be activated in young children by having them return to the exact same setting (i.e., same experimenter, same room, same set-up) involving highly distinct cues. To examine *which* of these overlapping features may be most responsible for the spontaneous retrieval, we repeated a basic experimental design (Krøjgaard et al., 2017) while manipulating a presumably central cue: whether 35- and 46-month-old children were met by the same or a new experimenter when returning to the lab for the test. Our findings replicate and extend those of previous studies. First, we documented that the children generally did have spontaneous recollections, replicating previous findings (Krøjgaard et al., 2017; Sonne et al., 2019) and lending further credibility to the paradigm used and the idea that spontaneous recall develops at least around 35-months of age. Furthermore, we found that the manipulation of the experimenter as a cue had different effects depending on the age of the children. For the 35-month-olds, we found no differences between the two conditions, whereas for

Notice: This is the author's version of a work that was accepted for publication in *Psychological Research.* A definitive version was subsequently published in *Psychological Research*. DOI: 10.1007/s00426-020-01310-7

the 46-month-olds, the spontaneous recall surprisingly was enhanced in the Change of Experimenter Condition compared to the Same Experimenter Condition. The finding that the change of experimenter had an effect on the performance in the oldest age group is thought-provoking given that the experimenter herself was never present during the actual 2 min Test. The additional analyses offered some explanations as to why the 46-month-olds performed better when met by a new experimenter on the second visit.

The fact that the 46-month-olds produced more spontaneous utterances in the Change of Experimenter Condition tended to be accompanied by the children mentioning this change, and/or providing descriptions of what it was like at the previous visit (i.e., more unspecific scores), and not by providing additional utterances related to the target event as such. For some of the 46-month-olds, noticing the change of experimenter seemed to serve as a distinct memory cue, potentially leading to sustained attention to the laboratory setting to which they returned. Thus, the additional unspecific scores in this condition may potentially be explained by a desire to remember or confirm what actually happened during the previous visit. Importantly, given that the experimenter was never present during the two-minute waiting period, the increase in spontaneous utterances cannot be explained by a motivation to brief this unfamiliar person about what happened during this test.

For the youngest age group, we saw a different pattern of results. The lack of reliable differences between the two conditions in this age group potentially suggests that the experimenter as a cue was less important for this age group. Considering the results of the study by Goldenberg and Sandhofer (2013), this may not be surprising. Perhaps due to poorer source monitoring skills, the children in the youngest age group in our study did not explicitly mention the change in experimenter in the two-minute test. This accords with the results of Goldenberg and Sandhofer (2013), who also tested three-year-olds. Other studies have also documented difficulties in this age group in relation to source monitoring. For instance, O'Neill and Gopnik (1991) found that children around the age of three (but not children aged four or five) encountered problems with this aspect (see also Gopnik & Graf, 1988). This interpretation is also in line with the 35-month-olds' poor performance on the follow-up question aimed at investigating whether they noticed the change. However, for two reasons, the answers to the presented question (Q: 'So at the previous visit you were presented with something inside that box [experimenter points at the target box]. Did *I* show it to you?') should probably be treated with some caution. First, we know that asking young children yes/no questions has inherent potential pitfalls (e.g., regarding accuracy) (for an overview see Pipe & Salmon, 2009). Second, these questions may have caused a potentially problematic asymmetry across the two experimental conditions: To answer correctly, the children in the Change of Experimenter condition had to say 'no', whereas those in the Same Experimenter Condition had to provide a 'yes' response. Knowing that children have a bias towards saying 'yes' to yes/no questions (Peterson, Dowden, & Tobin, 1999) – a bias that may have been prominent in the obtained results – we prefer to refrain from making firm conclusions regarding the results obtained by this question. A forced choice procedure would probably have been a better option and is therefore recommended for future studies.

Another explanation could be that we did not obtain significant differences between the two conditions for the youngest age group simply due to lack of statistical power. However, given that we observed a numerical difference in the opposite direction for this age group, this explanation seems unlikely. This drop in performance in the youngest age group, although not significant, accords with our original hypotheses as well as the results from studies focusing on infant memory (e.g., Learmonth et al., 2005), which document that memory during the first years of life is highly context dependent.

Notice: This is the author's version of a work that was accepted for publication in *Psychological Research.* A definitive version was subsequently published in *Psychological Research*. DOI: 10.1007/s00426-020-01310-7

Involuntary memories have been identified as fundamental in directing our behavior in novel contexts (Rasmussen & Berntsen, 2009). Returning to a specific location may be able to evoke a memory of a previous visit at this location, thus allowing one to identify potential differences between the two situations, and hence making it easier to predict what will happen next (e.g., Hintzman, 2011). In this case, returning to the same setting served as a cue, reminding the children that the experiment at the previous visit was different, thereby enabling them to add this knowledge to a mental representation of this event for potential future use. Similarly, as suggested by Wahlheim and Zacks (2019), noticing changes between a current and a previously experienced event is key to update our mental representations of the event. If such updating fails, this can have detrimental effects for our ability to predict what will happen later. However, if changes are remembered by the perceiver, this can lead to enhanced memory and event knowledge for events that are similar in kind. From our results, it would appear that 46-month-old children are more likely than 35-month-olds to engage in such cognitive processes. This line of thought corresponds with Visual Paired-Comparison data showing that 40-month-olds demonstrated prolonged visual attention to a video of a previously experienced lab visit with a new experimenter, relative to a video of the same lab visit with the original experimenter (Kingo, Staugaard, & Krøjgaard, 2014).

The results of our study raise some important questions for future investigation. When examining the importance of cues, different age groups should be examined. In the only previous attempt to examine the potential role of distinct cues for spontaneous recall using the same basic design, Sonne et al. (2019) showed that changing the context at test had a detrimental (although not significant) effect on spontaneous recall in 35-month-olds. The results obtained in the present study add to Sonne et al.'s findings in at least two respects. First, the results obtained here provide evidence that memory cues may have different effects in different age groups. Because the study by Sonne and colleagues (2019) only tested 35-month-olds, it is unclear whether the same pattern

Notice: This is the author's version of a work that was accepted for publication in *Psychological Research.* A definitive version was subsequently published in *Psychological Research*. DOI: 10.1007/s00426-020-01310-7

would hold for different age groups, or whether an interaction effect would have appeared similar to that in the present study. Second, whereas changing the context for recall (i.e., the room) induced concrete and tangible contextual differences between the baseline and the test for the two-minute waiting periods in Sonne et al.'s (2019) study, this was not true of the present study, because the experimenter was never present during these waiting periods. Thus, even for young children, we may have to think of memory cues for spontaneous recall in a slightly broader way than merely considering cues as only directly present in the here and now. Interestingly, however, the results of both the present study and that of Sonne et al. (2019) indicate that even when changing some contextual cues as well as some social cues, spontaneous recollections are quite robust.

Furthermore, the results lend further credibility to the paradigm used for activating spontaneous memories in children. Until now, we have examined the potential impact of changing the context for retrieval, as well as changing the experimenter. The finding that children consistently experience spontaneous recollections regardless of changing the location (Sonne et al., 2019) or the experimenter (the present study), seems to imply that it is a robust effect. Findings on spontaneously arising autobiographical memories in adults suggest that the memory cue typically is a central detail of the remembered event, defined as a detail that cannot not be replaced by a realistic alternative from the same category without changing the target of the subject's attention in the remembered event (Berntsen, 1998, 2009). Thus, central cues are distinctive for what the subject paid attention to in the specific, remembered situation. Following these findings, the presence of the distinct boxes may be the crucial cues responsible for having spontaneous memories. This suggestion could be tested empirically by at least two different approaches: First, if the boxes are crucial cues, then we should expect that the children would look reliably longer at the 'target box' (i.e., the box containing the event they were shown at T_1) relative to the foil box (the box containing the event they did not see at T_1) prior to producing spontaneous recollections at test. Second, if children returned to a location in which the boxes had been removed, it should reliably reduce the number of spontaneous recollections. We are currently pursuing these possibilities in our lab. Finally, the potential interaction of two or more of these cues still needs to be examined. Based on the available existing evidence we would suspect that combinations involving the presence of the boxes would be the most potent.

Although the basic design employed here appears to be a versatile paradigm for examining spontaneous recall in a controlled setting, a clear limitation worth mentioning is that the current design only allows us to capture what the children are stating verbally. That is, we do not have access to the spontaneous recollections that the children may have in their minds yet do not explicitly express. One might argue that this is a limitation present in any study using children's verbal responses as data source. However, in our view this limitation is likely to be especially prominent under the testing circumstances used in the present design. Because we explicitly instructed the parents not to interfere or ask questions, some of the children may have been left with the impression that they ought to remain silent during the 2-minute waiting periods. Anecdotal evidence lends support to this impression, as many of the children were whispering instead of speaking out normally while waiting with their parents during these 2-minute waiting periods. Thus, we find it likely that some of the children may have experienced spontaneous recollections but did not feel like verbalizing them due to the somewhat "artificial" testing conditions. Provided that our reasoning is correct, the present way of operationalizing spontaneous recall may in fact be somewhat conservative. Consequently, when using this paradigm there is the inherent risk that we may underestimate the prevalence of spontaneous recall in young children. Moreover, some children may be more likely to explicitly express what they are thinking, and as such the current design may only allow us to capture the spontaneous memories of these individuals. The nature of the unfamiliar lab setting may add to this potential problem. It would be interesting to investigate possible individual differences in children's spontaneous recall. Relatedly, when focusing on the importance of different cues, the current design only enables us to capture potential differences caused by external cues. We do not have access to any potential self-cueing strategies that the children may be employing, and are therefore unable to manipulate this aspect of the possible associative mechanisms involved in spontaneous recall. Again, this is an inherent limitation of insisting on a third-person perspective, and unfortunately we see no immediate way to circumvent this problem.

Despite these limitations, we believe that the present study provides important findings regarding the robustness of, and the importance of cues for, spontaneous recall in young children. When activating these memories, it seems that not only the overlap or distinctiveness of cues is sufficient for consideration. Distinct cues can have varied effects for different age groups, and depending on the type of cue, a lack of complete overlap between encoding and testing may not always have negative consequences, as long as other distinctive cues are present.

Disclosure statement:

The authors report no conflict of interest.

References

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

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

Berntsen, D. (2012). Spontaneous recollections: Involuntary autobiographical memories are a basic mode of remembering. In D. Berntsen & D. C. Rubin (Eds.), *Understanding autobiographical memory: Theories and approaches* (pp. 290-310). Cambridge: Cambridge University Press.

Berntsen, D., Staugaard, S. R., & Sørensen, L. M. T. (2013). Why am I remembering this now? Predicting the occurrence of involuntary (spontaneous) episodic memories. *Journal of Experimental Psychology: General, 142*, 426-444. http://dx.doi.org/10.1037/a0029128

Goldenberg, E. R., & Sandhofer, C. M. (2013). Who is she? Changes in the person context affect categorization. *Frontiers in Psychology, 4*, 1-5. https://doi.org/10.3389/fpsyg.2013.00745

Gopnik, A., & Graf, P. (1988). Knowing how you know: Young children's ability to identify and remember the sources of their beliefs. *Child Development, 59*, 98-110. doi: 10.2307/1130499

Hanna, E., & Meltzoff, A. N. (1993). Peer imitation by toddlers in laboratory, home, and day-care contexts: Implications for social learning and memory. *Developmental Psychology*, *29*(4), 701-710. http://psycnet.apa.org/doi/10.1037/0012-1649.29.4.701

Hayne, H. (2004). Infant memory development: Implications for childhood amnesia. *Developmental Review, 24*, 33-73. http://dx.doi.org/10.1016/j.dr.2003.09.007

Hintzman, D. L. (2011). Research strategy in the study of memory: Fads, fallacies, and the search for the "coordinates of truth". *Perspectives on Psychological Science, 6,* 253-271. https://doi.org/10.1177/1745691611406924

Hunt, R. R., & Smith, R. E. (1996). Accessing the particular from the general: The power of distinctiveness in the context of organization. *Memory & Cognition, 24,* 217-225.

Johnson, M. H. (2005). *Developmental cognitive neuroscience* (2nd ed.). Oxford: Blackwell Publishing.

Kingo, O.S., Staugaard, S.R., & Krøjgaard, P. (2014). Three-year-olds' memory for a person met only once at the age of 12 months: Very long-term memory revealed by a late-manifesting novelty preference. *Consciousness and Cognition, 24,* 49-56, doi: 10.1016/j.concog.2013.12.011.

Krøjgaard, P., Kingo, O. S., Dahl, J. J., & Berntsen, D. (2014). "That one makes things small": Experimentally induced spontaneous memories in 3.5-year-olds. *Consciousness and Cognition, 30,* 24-35. doi: 10.1016/j.concog.2014.07.017

Krøjgaard, P., Kingo, O. S., Jensen, T. S., & Berntsen, D. (2017). By-passing strategic retrieval: Experimentally induced spontaneous episodic memories in 35- and 46-month-old children. *Consciousness and Cognition, 55*, 91-105. doi: 10.1016/j.concog.2017.08.001

Learmonth, A. E., Lamberth, R., & Rovee-Collier, C. (2005). The social context of imitation in infancy. *J. Experimental Child Psychology, 91,* 297-314. doi: 10.1016/j.jecp.2005.02.001

Levine, B., Svoboda, E., Hay, J.F., Winocur, G, & Moscovitch, M. (2002). Ageing and autobiographical memory: Dissociating episodic from semantic retrieval. *Psychology and Aging, 17,* 677-689. doi: 10.1037//0882-7974.17.4.677

Nairne, J. S. (2002). The myth of the encoding-retrieval match. *Memory, 10,* 389-395. doi:

10.1080/09658210244000216

Nelson, K., & Ross, G. (1980). The generalities and specifics of long-term memory in infants and young children. *New Directions for Child and Adolescent Development, 10,* 87-101.

O'Neill, D. K., & Gopnik, A. (1991). Young children's ability to identify the sources of their beliefs. *Developmental Psychology, 27*(3), 390-397. http://dx.doi.org/10.1037/0012-1649.27.3.390

Peterson, C., Dowden, C., & Tobin, J. (1999). Interviewing preschoolers: comparisons of yes/no and wh- questions. *Law and Human Behavior, 5,* 539-555.

Pipe, M., & Salmon, K. (2009). Memory development and the forensic context. In M. L. Courage & N. Cowan (Eds.), *The development of memory in infancy and childhood.* (2nd ed., pp. 241-282). Hove and New York: Psychology Press.

Rasmussen, A. S., & Berntsen, D. (2009). The possible functions of involuntary autobiographical memories. *Applied Cognitive Psychology, 23,* 1137-1152. https://doi.org/10.1002/acp.1615

Reese, E. (1999). What children say when they talk about the past. *Narrative Inquiry, 9,* 215-241. http://dx.doi.org/10.1075/ni.9.2.02ree

Sonne, T., Kingo, O. S., Berntsen, D., & Krøjgaard, P. (2019). Thirty-five-month-old children have spontaneous memories despite change of context for retrieval. *Memory, 27,* 38-48. doi: 10.1080/09658211.2017.1363243

Sonne, T., Kingo, O. S., & Krøjgaard, P. (2016). Empty looks or paying attention? Exploring infants' visual behavior during encoding of an elicited imitation task. *Infancy*, *21*, 728-750. https://doi.org/10.1111/infa.12141

Todd, C.M., & Perlmutter, M. (1980). Reality recalled by preschool children. *New Directions for Child and Adolescent Development, 10,* 69-85.

Tulving, E., & Thomson, D. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80,* 352-373.

Wahlheim, C. N., & Zacks, J. M. (2019). Memory guides the processing of event changes for older and younger adults. *Journal of Experimental Psychology: General, 148,* 30-50.

Notice: This is the author's version of a work that was accepted for publication in *Psychological Research.* A definitive version was subsequently published in *Psychological Research*. DOI: 10.1007/s00426-020-01310-7

Table 1: Word List in alphabetical order for the two events, including the list of 'Unspecific' words indicating that the child had been there before but without referring to specific event details. Please note that synonyms and conjugations of the words presented in the table were also included.

Teddy Condition	Game Condition	Unspecific
Alfred	Ball	'Always'
Animal	Blue	'Have been'
Button	Bowling	'Again'
Clap	Bucket	'Last time'
Dog	Game	'Key'
Elephant	Green	'Unlock'
Elly	Hit	'New/other girl/lady
Ear	Medal	'Toys'
Hug	Pin	
Lives	Play	
Music	Prize	
Push	Red	
Sing	Roll	
Take/bring out or get	Throw	
Teddy	Turn over	
Turn on	Win	
Wiggle	Yellow	

Notice: This is the author's version of a work that was accepted for publication in *Psychological Research.* A definitive version was subsequently published in *Psychological Research*. DOI: 10.1007/s00426-020-01310-7

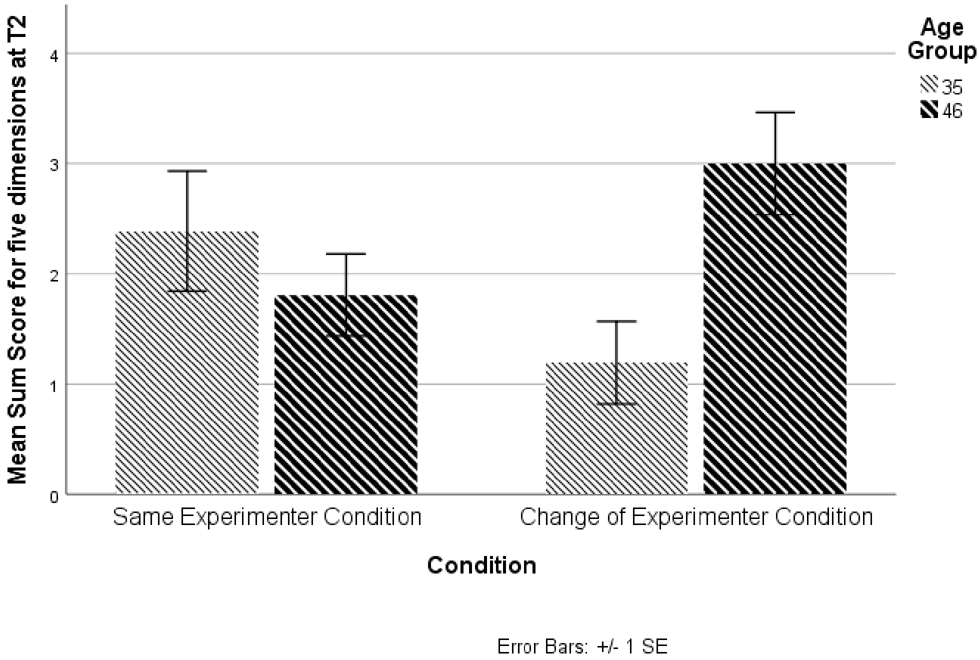
Table 2 displays the descriptive statistics and main effects based on the Word List and Coding Scheme. Based on previous studies using the same design (e.g., Krøjgaard et al., 2017; Sonne et al., 2019) the congruent hits include both condition-specific hits (e.g., "teddy", "elephant") and unspecific hits (e.g., "last time", "again").

	Same Experimenter Condition					
35-month-olds (n = 31)	**Descriptives**		**One-sample t-tests**			
Measure	*M*	*SD*	*t*	*df*	*p*	*r*
Word List	1.10	1.599	3.82	30	= .001	.57
Coding Scheme						
Language	0.74	0.97	4.28	30	<.001	.62
Gesture	0.32	0.48	3.78	30	=.001	.57
Reliving	0.74	0.97	4.28	30	<.001	.62
Action	0.35	0.61	3.25	30	=.003	.51
Spatial	0.23	0.50	2.53	30	=.017	.42
Social	0.03	0.18	1.00	30	=.325	.18
SUM Score (Coding Scheme)	2.39	3.03	4.39	30	<.001	.63
46-month-olds (n = 31)	**Descriptives**		**One-sample t-tests**			
Measure	*M*	*SD*	*t*	*df*	*p*	*r*
Word List	0.97	1.52	3.55	30	.001	.54
Coding Scheme						
Language	0.55	0.62	4.89	30	<.001	.67
Gesture	0.32	0.48	3.78	30	=.001	.57
Reliving	0.55	0.62	4.89	30	<.001	.67
Action	0.19	0.40	2.68	30	=.012	.44
Spatial	0.19	0.40	2.68	30	=.012	.44
Social	0.00	0.00	-	30	-	-
SUM Score (Coding Scheme)	1.81	2.07	4.85	30	<.001	.66

	Change of Experimenter Condition					
35-month-olds (n = 31)	**Descriptives**		**One-sample t-tests**			
Measure	*M*	*SD*	*t*	*df*	*P*	*r*
Word List	0.55	1.179	2.59	30	=.015	.43
Coding Scheme						
Language	0.42	0.77	3.05	30	=.005	.49
Gesture	0.13	0.34	2.11	30	=.043	.36
Reliving	0.45	0.77	3.28	30	=.003	.51
Action	0.13	0.34	2.11	30	=.043	.36
Spatial	0.06	0.25	1.44	30	=.161	.25
Social	0.00	0.00	-	30	-	-
SUM Score (Coding Scheme)	1.19	2.09	3.18	30	=.003	.50
46-month-olds (n = 31)	**Descriptives**		**One-sample t-tests**			
Measure	*M*	*SD*	*t*	*df*	*p*	*r*
Word List	1.19	1.33	5.01	30	<.001	.67
Coding Scheme						
Language	1	0.93	5.98	30	<.001	.74
Gesture	0.42	0.62	3.76	30	=.001	.57
Reliving	1.10	1.01	6.04	30	<.001	.74
Action	0.29	0.46	3.50	30	=.001	.54
Spatial	0.19	0.40	2.68	30	=.012	.44
Social	0.03	0.18	1.00	30	=.325	.18
SUM Score (Coding Scheme)	3	2.58	6.47	30	<.001	.76

Notice: This is the author's version of a work that was accepted for publication in *Psychological Research.* A definitive version was subsequently published in *Psychological Research*. DOI: 10.1007/s00426-020-01310-7

Figure 1



Notice: This is the author's version of a work that was accepted for publication in *Psychological Research.* A definitive version was subsequently published in *Psychological Research*. DOI: 10.1007/s00426-020-01310-7

Figure caption

Figure 1: Bar Chart displaying the mean number of sum scores from five dimensions from the Coding Scheme (social dimension excluded), obtained at T_2 across age groups and conditions.