Notice: This is the author's version of a work that was accepted for publication in *Consciousness and Cognition.* A definitive version was subsequently published in *Consciousness and Cognition, 55, 91-105.* DOI: 10.1016/j.concog.2017.08.001

By-passing strategic retrieval: Experimentally induced spontaneous episodic memories in 35- and 46-month-old children

Peter Krøjgaard*
Osman S. Kingo*
Toril S. Jensen*
Dorthe Berntsen*

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

Corresponding Author:

Peter Krøjgaard, Center on Autobiographical Memory Research, Department of Psychology and Behavioural Sciences, Aarhus University, Bartholins Allé 9, Bld. 1350, 8000 Aarhus C, Denmark, Mail: peter@psy.au.dk, Phone: 0045 87165861.

Co-authors:

Osman S. Kingo, Center on Autobiographical Memory Research, Department of Psychology and Behavioural Sciences, Aarhus University, Bartholins Allé 9, Bld. 1350, 8000 Aarhus C, Denmark, Mail: osman@psy.au.dk, Phone: 0045 87165862.

Toril S. Jensen, Center on Autobiographical Memory Research, Department of Psychology and Behavioural Sciences, Aarhus University, Bartholins Allé 9, Bld. 1350, 8000 Aarhus C, Denmark, Mail: torilsj@psy.au.dk, Phone: 0045 87165884.

Dorthe Berntsen, Center on Autobiographical Memory Research, Department of Psychology and Behavioural Sciences, Aarhus University, Bartholins Allé 9, Bld. 1350, 8000 Aarhus C, Denmark, Mail: dorthe@psy.au.dk, Phone: 0045 87165868.

Notice: This is the author's version of a work that was accepted for publication in *Consciousness and Cognition.* A definitive version was subsequently published in *Consciousness and Cognition, 55, 91-105.* DOI: 10.1016/j.concog.2017.08.001

Abstract

Most parents have experienced their preschool child having spontaneous episodic memories, that is, verbally reported memories of past events that come to the child almost out of the blue. Until recently such memories had only been observed outside the lab. By means of a new paradigm we report experimentally induced spontaneous memories of a unique event experienced one week earlier in 35- and 46-month-old children (N=110). At the first visit, half of the children experienced a Teddy event and the other half experienced a Game event. At the second visit the children's spontaneous utterances were recorded while waiting. The results revealed that the children talked spontaneously about the unique event experienced previously. Age showed no systematic effect on spontaneous episodic recollection, but there was a clear effect of age on subsequent control questions requiring strategic retrieval. The results support the idea of involuntary episodic remembering being a developmentally early achievement.

Keywords: spontaneous memories; episodic memories; young children; retrieval; methodology

Notice: This is the author's version of a work that was accepted for publication in *Consciousness and Cognition.* A definitive version was subsequently published in *Consciousness and Cognition, 55, 91-105.* DOI: 10.1016/j.concog.2017.08.001

By-passing strategic retrieval: Experimentally induced spontaneous episodic memories in 35- and 46-month-old children

1.0 Introduction

The predominant method used when examining verbally reported memories in children is to simply ask the children about their previous experiences (Hayne, Scarf, & Imuta, 2015). When asked directly about previously experienced events, young children often have a hard time retrieving and talking about specific episodes from their past (e.g., Dahl, Kingo, & Krøjgaard, 2015; Simcock & Hayne, 2002). In order to respond to such questions, children will have to engage in a deliberate and strategic retrieval process. Deliberate and strategic retrieval requires executive control and involves the use of frontal lobes, which matures relatively late in the ontogenesis (e.g., Johnson, 2005). Thus, for young children, the process involved and required when responding to direct questions about past events, is by itself a cognitively demanding achievement which, all things equal, may have detrimental effects on the children's memory performance.

However, strategic recall is not the only way in which memories of past events come to mind. Evidence from healthy adults has shown that memories frequently come to mind involuntarily, or spontaneously, that is, without any deliberate attempts to recall the episode, but often in response to concrete environmental cues (Berntsen, 1996, 2009). Involuntary episodic memories differ from voluntary episodic memories by not involving neural activity in prefrontal areas associated with retrieval effort (Hall et al., 2014) and by involving shorter retrieval time (e.g., Berntsen, Staugaard, & Sørensen, 2013; Schlagman & Kvavilashvili, 2008), both consistent with the idea of relatively effortless retrieval. At the same time, such involuntary memories appear to be highly cue dependent (Berntsen, 1996, 2009). In addition

Notice: This is the author's version of a work that was accepted for publication in *Consciousness and Cognition.* A definitive version was subsequently published in *Consciousness and Cognition, 55, 91-105.* DOI: 10.1016/j.concog.2017.08.001

involuntary memories are more frequently about specific episodes and have been found to have more emotional impact, and some studies find them to be more vivid than voluntary memories (for a review, see Berntsen, 2010). The notion of spontaneous memories as used here is adopted from the research on involuntary memories in adults.

We define spontaneous memories as (i) verbally produced, (ii) socially unprompted, and (iii) environmentally cued (Krøjgaard, Kingo, Dahl, & Berntsen, 2014). 'Socially unprompted' means that the memories do not occur as a result of prompts of any kind, as for instance explicit or implicit questions or demand characteristics directed at the child. Because spontaneous retrieval is based primarily on associative mechanisms and hence less dependent on executive functions and mature frontal lobes, (a) spontaneous retrieval is assumed to be less cognitively demanding than strategic retrieval, and, following the same logic, (b) spontaneous memories may be more prevalent in children than in adults, and (c) is likely to be developmentally earlier than strategic remembering (Berntsen, 2009, 2012). In the following we briefly present the existing evidence on spontaneous memories in children (for a more detailed review, see Krøjgaard et al., 2014).

1.1. Spontaneous memories in children

We recently reported experimental evidence of spontaneous recollections of a repeated, non-specific event in 46 months old children (Krøjgaard et al., 2014). However, no systematic research has been conducted on children's spontaneous recollections of unique episodes in their past. The literature is also lacking systematic examination of the effects of age on spontaneous versus strategic episodic memories.

Although previous work is scarce, at times spontaneous memories as defined here have appeared in studies examining children's memory in general. Spontaneous memories have for instance been observed in a number of studies using unstructured or semi-structured

Notice: This is the author's version of a work that was accepted for publication in *Consciousness and Cognition.* A definitive version was subsequently published in *Consciousness and Cognition, 55, 91-105.* DOI: 10.1016/j.concog.2017.08.001

methodologies, usually diary studies (Ashmead & Perlmutter, 1980; Hudson, 1990; Nelson, 1989; Nelson & Ross, 1980; Reese, 1999; Todd & Perlmutter, 1980). As an illustrative example of a spontaneous memory, Todd and Perlmutter (1980, p. 82) reported the following:

> […] the parents of one of the three-year-olds reported that while watching a commercial that displayed a bottle of honey, the child said she had liked the chocolate stuff her mother used to give her but not the yellow. The mother felt this could only refer to the chocolate syrup and honey she used to put in the child's bottle when she was a baby.

A number of experimentally based studies have collected children's 'spontaneous' verbal references to their previous visits to a lab, while their non-verbal memories were tested by means of the elicited imitation paradigm (Bauer, Kroupina, Schwade, Dropik, & Wewerka, 1998; Bauer, van Abbema, Wiebe, Cary, Phill, & Burch, 2004; Bauer, Wenner, & Kroupina, 2002; Bauer & Wewerka, 1995). To illustrate, in some of the studies by Bauer and colleagues the children 'spontaneously' reported task-relevant material from the previous lab visits during the first of two delayed tests (Bauer et al., 1998; Bauer et al., 2004; Bauer et al., 2002; Bauer & Wewerka, 1995). However, because the setting involved encouragements to remember the previous task, as the children were explicitly asked to recall the motor actions of the events (i.e., "You can make a windmill with this stuff. Show me how to make a windmill"), these verbal productions were prompted and would therefore not fulfill the criteria used here for spontaneous memories.

Krøjgaard et al. (2014) presented evidence of experimentally induced spontaneous memories of a repeated event. Forty-six-month-old children were brought back to the lab where they had previously experienced memorable events, and while waiting in front of the

Notice: This is the author's version of a work that was accepted for publication in *Consciousness and Cognition.* A definitive version was subsequently published in *Consciousness and Cognition, 55, 91-105.* DOI: 10.1016/j.concog.2017.08.001

props for the experimenter to return from a "phone call", their possible spontaneous utterances were recorded. The rationale behind being able to induce spontaneous memories from specifically these previous encounters in the lab (and not any other experiences) was based on two assumptions: First, by bringing the children back to the exact same lab, with the same props, the same experimenter etc., there was a substantial overlap between the recall situation and the to-be-remembered episode which is known to facilitate memory (Hayne, 2004; Newcombe, Lloyd, & Ratliff, 2007). Second, the cues involved were not only overlapping; they were highly distinct. This strategy was consistent with the notion of cue overload stating that "The probability of recalling an item declines with the number of items subsumed by its functional retrieval cue" (Watkins & Watkins, 1975, p. 442). Thus, the likelihood of a cue providing access to a given target memory depends on the extent to which this cue is uniquely associated with the target. Its strength declines to the extent it is associated with other memories as well. This principle has shown its relevance in relation to involuntary episodic memories in adults (e.g., Berntsen et al., 2013) and in non-human primates (Martin-Ordas, Berntsen, & Call, 2013). Based on this principle, the most important factor is not the number or range of overlapping features between a present context and a remembered event, but rather whether the available cues referred *specifically* to the to-be-remembered episode, and not to other episodes. The recordings from the waiting period revealed that the 46-month-olds spontaneously produced significantly more mnemonic material relative to an age-matched control group (Krøjgaard et al., 2014).

Whereas the study by Krøjgaard et al. (2014) demonstrated that spontaneous memories could be induced experimentally, the study had a number of limitations: First, only 46-month-olds were examined; hence we do not know whether spontaneous memories could be induced in even younger children. Second, the props were fully visible during the test,

Notice: This is the author's version of a work that was accepted for publication in *Consciousness and Cognition.* A definitive version was subsequently published in *Consciousness and Cognition, 55, 91-105.* DOI: 10.1016/j.concog.2017.08.001

whereby online reasoning could be a confounder. Third, because the children in the experimental group had visited the lab three times before, the to-be-remembered event was not unique and therefore did not fulfill a commonly used criterion for episodic memory. In addition, the multiple previous visits may have made the children in the experimental group more comfortable during the test relative to controls that had never been in the lab before. Fourth, the previous visits concerned memory tasks. Thus, the possibility remains that although the children were never asked or prompted while waiting in front of the props, some of the children in the experimental group may have guessed that this study was about memory too.

In order to pursue the investigation of experimentally induced spontaneous recall further, we here tested the children's memories for a single distinct event, contrasted age groups and improved the experimental procedure used by Krøjgaard et al. (2014).

1.2 The present study

In the present study we attempted to induce spontaneous memories in 35- and 46-month-old children, and relate the results from spontaneous recall to the results from strategic and deliberate recall obtained in subsequent control questions. Thus, when examining different age groups, we may see a change in the pattern of results across age and mode of retrieval. Forty-six-month-old children were chosen, because this was the age of the children participating in the study we set out to improve (Krøjgaard et al., 2014). The 35-month-old children were chosen in order to see, if spontaneous recall could be induced in substantially younger children, that is, in children below the age of 3.

All children visited the lab twice with a one-week retention interval. At the first visit the children were shown one out of two possible interesting and highly distinct events: either a Teddy Event involving mechanical teddies that could sing, clap and wiggle their ears; or a

Notice: This is the author's version of a work that was accepted for publication in *Consciousness and Cognition.* A definitive version was subsequently published in *Consciousness and Cognition, 55, 91-105.* DOI: 10.1016/j.concog.2017.08.001

Game Event involving two funny games in which the children won medals. The props for the events resided in two differently looking, opaque and locked boxes. One week later the children returned to the lab, and while waiting for the experimenter to return to the test room, the children's possible spontaneous utterances were recorded. In addition, after the recollection of spontaneous utterances, the children were asked control questions requiring strategic recall.

We expected the 46-month-old children to produce more spontaneous utterances concerning the event they had been exposed to at the first visit, relative to the event they had not seen before. Since no previous studies have examined experimentally induced spontaneous recall in 35-month-old children, our expectations could only be speculative. Nonetheless, since spontaneous recall is less demanding than strategic recall, we expected the 35-month-old to also demonstrate spontaneous recollections of the staged event, potentially at the same level as the older children. In contrast, because the control questions would require strategic and deliberate recall, which involves prefrontal processes to monitor the search, we expected the 35-month-old children to have more problems with the control questions relative to their older peers.

2.0 The Experiment

2.1 Participants

A total of 110 children participated in the study. Sixty 35-month-olds (22 female, M_{age} = 35.62 months, *SD* = .51; range: 35.00- 37.73 months) and fifty 46-month-olds (30 female, M_{age} = 46.75 months, *SD* = 1.15; range: 45.13-49.63 months). In each age group the children were randomly assigned to one of two conditions, either the Teddy Condition or the Game Condition. Three additional children were tested but not included in the analyses (experimental error: 2; technical error: 1). The children were predominantly Scandinavian

Notice: This is the author's version of a work that was accepted for publication in *Consciousness and Cognition.* A definitive version was subsequently published in *Consciousness and Cognition, 55, 91-105.* DOI: 10.1016/j.concog.2017.08.001

Caucasian, living in families with middle to higher SES. All the children were healthy and full-term and were recruited via registers from the National Board of Health. The children received a small gift for their participation.

2.2 Materials

The children were tested in a sparsely furnished 16 m^2 room. Along one side, two boxes were placed right next to each other (see Fig. 1): To the left a red metal box with two doors measuring (W x H x D: 119 cm x 63cm x 40 cm), and to the right a brown wooden box with a top lid (W x H x D: 127 cm x 63cm x 51 cm). Both boxes were opaque and had locks. Along the opposite side of the room was a table (W x D: 180 cm x 90 cm) with chairs on the side from where there was a clear view to both boxes. Each box contained specific props to provide a unique and distinct experience dependent on whether the child had been allocated to the Teddy Condition (the red metal box) or the Game Condition (the brown wooden box). A camcorder recorded all sessions for later scoring.

2.2.1 Materials for the Teddy Condition.

The red metal box contained two distinct mechanical toy teddies: A blue elephant named Elly (W x H x D: 20 cm x 24 cm x 28 cm), who could sing a song while wiggling its ears when activated; a light-brown dog named Alfred (W x H x D: 20 cm x 26 cm x 20 cm), who could sing a song while wiggling its ears and clapping its paws when activated. None of these animated teddies were known to the children before the experiment; nor is this type of animated toys easily purchasable in the children's country. Both teddies were imported for the purpose of the experiment.

2.2.2 Materials for the Game Condition.

The brown wooden box contained props for two games: A throwing game, a bowling game, and a large green cloth to be put on the floor before playing. The materials for the

Notice: This is the author's version of a work that was accepted for publication in *Consciousness and Cognition.* A definitive version was subsequently published in *Consciousness and Cognition, 55, 91-105.* DOI: 10.1016/j.concog.2017.08.001

throwing game were three colored (blue, red, and green) plastic buckets glued to a plate and three soft balls. For the bowling game two plastic balls and ten plastic pins were used. As with the teddy condition, these games were unlikely to be known in advance to the children. The throwing game was constructed for the experiment, and bowling is an atypical game for young children.

2.3 Procedure

During recruitment, the parents were thoroughly informed that the study concerned spontaneous recall and they were instructed not to talk about the study and not in any way let the children know that the study was about memory. In case the child should ask about the visits, the parents were instructed to reply "I don't know". Similar precautions were made with regards to the children by ensuring that none of them had participated previously in any studies in our lab or currently participated in parallel studies. The purpose was to rule out that the children would guess or be left with the impression that the study was about memory. The children visited the lab twice with a one-week retention interval (± 1 day). One experimenter (the third author) recruited the participants and conducted all experiments.

2.3.1 T_1 – Encoding.

At the first visit (T_1), the procedure involved two steps: (1) A two-minute baseline measure, and (2) a six-minute encoding session involving the Condition to which the child was allocated (see Fig. 2).

When picked up in the waiting room, the parent was handed a written note as a reminder that a camera would be running when entering the lab. The note also reminded the parent not to initiate any conversation and only to respond briefly and in confirmatory fashion, but not to follow up on any statements from the child during the baseline assessment. In the lab the child and the parent were seated next to each other facing the two locked boxes

Notice: This is the author's version of a work that was accepted for publication in *Consciousness and Cognition.* A definitive version was subsequently published in *Consciousness and Cognition, 55, 91-105.* DOI: 10.1016/j.concog.2017.08.001

which were both clear in view. The experimenter then stated: "I've got to do something, but I will be right back." She left the lab, closed the door, and the child and the parent were left alone in front of the two boxes for exactly two minutes. The sequence was recorded for later scoring and any verbal utterances produced by the child served as a baseline measure (see below).

When the two minutes had passed, the experimenter returned. In order to assess the productive vocabulary of the children, the parent was handed a Danish version of the MacArthur Communicative Development Inventory: Words and Sentences (CDI) and asked to return it at the second visit.[1] At the second step (the condition specific encoding), each child was shown one, and only one, of the events: Children in the Teddy Condition took part in the Teddy event, whereas the children in the Game Condition took part in the Game event. Both events lasted approximately six minutes. The purpose of the encoding sequences was to show and engage each child in a highly distinct and enjoyable event. Importantly, at no time during the encoding session were the children given the impression that they had to remember anything; the word 'remember' (or any related words) did not appear in the two distinct event demonstrations. These precautions were made in order to rule out the possibility that the children would feel socially prompted to remember anything when returning for the test a week later.

2.3.1.1 The Teddy event. The experimenter began asking the child whether he or she could clap and wiggle their ears, and the child was encouraged to demonstrate the ability. Then the experimenter referred to the red metal box and opened it with a key. The experimenter took out the two teddies and told the child that these were very special teddies

[1] We failed to collect CDI reports from 1 out of the fifty 46-month-olds and from 2 out of the sixty 35-month-olds.

Notice: This is the author's version of a work that was accepted for publication in *Consciousness and Cognition.* A definitive version was subsequently published in *Consciousness and Cognition, 55, 91-105.* DOI: 10.1016/j.concog.2017.08.001

that could sing, clap, and wiggle their ears. The two teddies were then demonstrated. The child was asked about which animals the teddies were, and the names of the teddies (Elly the elephant, and Alfred the dog) were given. The child was encouraged to say the animals and the names out loud and praised if doing so. Each mechanical teddy was 'performing', and the child was encouraged to activate the teddies and praised for any engagement. The experimenter then encouraged the child to help out putting them back into the red metal cupboard after which it was locked. The experimenter ended the session by saying: "That was it for today. I'll be seeing you again next week!" In case questions were asked about the next meeting, the experimenter just replied: "I don't know. We'll see."

2.3.1.2 The Game event. The experimenter began by referring to the brown wooden box and opened it with a key. The experimenter unpacked the contents and told the child that they were about to play two funny games, and that the child could win medals. The green cloth was placed on the floor, and while informing the child, the experimenter brought the three buckets for the throwing game. The child was asked whether he or she knew the name of the items (buckets) and their colors. The child was encouraged to say the name and the colors out loud and praised if doing so. The experimenter then demonstrated the game: At close distance one had to 'throw' balls into the three buckets. No matter how the game went the child was praised and received a 'gold medal'. Then the game was put back into the wooden box and the props for the bowling game were arranged. Again the experimenter asked the child whether he or she knew the name of the game and the items involved, and the child was praised for saying the names out loud. Then the experimenter showed how to play bowling: one had three attempts to turn over the pins. No matter how the game went, the child received another gold medal and was praised. The child was encouraged to help putting the props back in the brown wooded box that subsequently was locked. The experimenter

Notice: This is the author's version of a work that was accepted for publication in *Consciousness and Cognition.* A definitive version was subsequently published in *Consciousness and Cognition, 55, 91-105.* DOI: 10.1016/j.concog.2017.08.001

ended the Game event in exactly the same manner as the Teddy Condition ("That was it for today." etc.).

2.3.2 T_2 – test.

The second visit (T_2) involved two steps: (1) a two minute test of spontaneous recall, and (2) control questions involving strategic retrieval. The two minute test was an exact replication of the baseline procedure administered at T_1. The only difference was that this time, the child had actually been in the lab once before and knew about the contents of one, but only one, of the boxes, dependent on condition.

When the two minutes had passed the experimenter returned to ask control questions. These questions served two purposes. First, they probed strategic recall by asking the child specific questions about what he or she had experienced at the first visit. Second, they served to make sure that the child did not erroneously produce details of the non-experienced event. The control questions were asked *after* the two-minute spontaneous memory test in order to conceal that the study was about memory until the test for possible spontaneous utterances was finalized. Following the Krøjgaard et al. (2014) study, the control questions used here were of a focused 'yes/no' type. For each of the two boxes (of which each child had only seen the contents of one), the child was asked two questions that required the children to search their memory to provide an answer. The order of the questions addressing the two events was balanced across subjects and conditions. The experimenter began by asking: "Remember last time you were here:

Q1: "Did you see what was in the red cupboard/brown wooden box?" [If the response was "no", or if a "yes" was accompanied by a description of the contents, question #2 was left out. If no response was given, then question #2 was asked]

Q2: "What was in the red cupboard/brown wooden box?"

Notice: This is the author's version of a work that was accepted for publication in *Consciousness and Cognition.* A definitive version was subsequently published in *Consciousness and Cognition, 55, 91-105.* DOI: 10.1016/j.concog.2017.08.001

The child was typically prompted up to twice to respond. If no response were given after the second prompt, then no further questions were asked.

2.4 Coding and data reduction

Two kinds of data were collected in the experiment: (i) spontaneous verbal responses produced while waiting for the experimenter to return at both T_1-baseline and T_2-test, (ii) responses to the control questions at T_2. We first consider possible spontaneous verbal responses.

2.4.1 Coding of spontaneous verbal responses.

Central for the coding of spontaneous verbal responses was the two videos of two-minute duration from each child: One video from T_1-baseline and one from T_2-test. Prior to the coding of the spontaneous verbal responses all video recordings were edited so they only contained the two-minute excerpts from when the children were waiting for the experimenter to return. Thus, the scorers were blind with regards to which condition a given child had been allocated to, as well as whether a given video originated from T_1-baseline or from T_2-test.

Based on the coding strategy employed in the Krøjgaard et al. (2014) study, all spontaneous verbal utterances produced by the children during the two-minute waiting time at both baseline and test were scored by two measures: By a Word List and by a Coding Scheme.

2.4.1.1 Scoring by Word List. From the edited recordings a primary scorer scored verbally communicated spontaneous memories related to the to-be-remembered event from the children based on a Word List (see Appendix A). The Word List consisted of words related specifically to each of the two events (The Teddy Event and The Game Event) and an additional Unspecific list referring to utterances indicating that the child had been in lab before, but without being specific about any of the two distinct events (e.g., right after being

Notice: This is the author's version of a work that was accepted for publication in *Consciousness and Cognition.* A definitive version was subsequently published in *Consciousness and Cognition, 55, 91-105.* DOI: 10.1016/j.concog.2017.08.001

left again at the T_2-test in the lab for the two-minute period, one child said: "She did that last time, too"). The Word List was fixed and was developed prior to the scoring. Any mentioning of a word from the Word List (see Appendix A) resulted in the score of either a "T" (Teddy), "G" (Game), or "U" (Unspecific) dependent of which of the three lists the word belonged to. Note that these distinctions were necessary as both scores were indicative of having been in the lab before. For each child sum scores were calculated.

2.4.1.2 Scoring by Coding Scheme. The Coding Scheme was developed in order to obtain an additional scoring of the possible spontaneous verbal utterances, in case there might be spontaneous verbal utterances indicative of memory for the events that would not be captured by the words from the fixed Word List. Each two-minute video sequence from both the T_1-baseline and the T_2-test "waiting" phase was divided into 12 time-slots of ten seconds' duration each. For each ten second time-slot the scorer had to reply to seven specific questions regarding possible mnemonic content in the video segments. The seven questions concerned the following seven dimensions: (1) *language*, (2) *gestures*, (3) *reliving*, (4) *action details*, (5) *object details*, (6) *spatial details*, and (7) *social details*. The Coding Scheme used here was a modified version of the Coding Scheme developed for the Krøjgaard et al. (2014) study, which was inspired by the coding of internal (episodic) details developed by Levine, Svoboda, Hay, Winocur and Moscovitch (2002) in order to distinguish between episodic and semantic components of memories of episodes. The exact wording of the seven questions was as follows:

1. Does the child by means of *language* refer to knowledge that originates from a previous visit (e.g., by expecting something to happen, or to procedures, actions, events, objects or details)?

Notice: This is the author's version of a work that was accepted for publication in *Consciousness and Cognition.* A definitive version was subsequently published in *Consciousness and Cognition, 55, 91-105.* DOI: 10.1016/j.concog.2017.08.001

2. Does the child by means of *gestures* refer to knowledge that originates from a previous visit (e.g., by expecting something to happen, or to procedures, actions, events, objects or details)?
3. Does the child's verbal and/or non-verbal behavior indicate that the child mentally *relives* parts of a previous visit (e.g., by imitating earlier actions or by expressing enthusiasm or joy originating from a previous visit)?
4. Does the child refer to specific *action details* from a previous visit (e.g., related to procedures, actions, events, or objects)?
5. Does the child by means of specific *object details* refer to knowledge that originates from a previous visit (e.g., related to the name, shape, color, size, texture, or function of an object)?
6. Does the child by means of specific *spatial details* refer to knowledge that originates from a previous visit (e.g., where things are, or where in the room something took place)?
7. Does the child by means of specific *social details* refer to knowledge that originates from a previous visit (e.g., by asking for specific persons/animals/teddies)?

For each ten second time-slot, the scorer answered the seven questions presented above. Whenever there was a hit, the scorer gave either a "T" (Teddy), "G" (Game), or "U" (Unspecific) dependent on whether the utterance referred specifically to one of the events demonstrated at encoding (Game or Teddy), or whether the utterance simply indicated that the child had been there before (Unspecific). For each question, each ten second time-slot could either give a single score or no score. For each of the seven questions, a sum score based on the results from the 12 ten second time-slots was calculated (range: 0-12). Thus, for each child seven sum scores were derived. Obviously, the possible responses to the seven

Notice: This is the author's version of a work that was accepted for publication in *Consciousness and Cognition.* A definitive version was subsequently published in *Consciousness and Cognition, 55, 91-105.* DOI: 10.1016/j.concog.2017.08.001

questions were not mutually exclusive. For example, one 46-month-old girl (Game cond.) happily said "I know what's inside, it's a game" which led to scorings in dimension 1, 2, 3, and 6, but not in 4, 5, and 7. To illustrate: Score on dimension 1 (*language*) was granted, since the statement was verbal. Score on dimension 2 (*gesture*) was granted as the girl pointed to the brown, wooden box, while producing the utterance. Score on dimension 3 (*relieving*) was given because the girl expressed enthusiasm and joy while having the spontaneous recollection (cf. the operationalization of dimension 3). And finally a hit on dimension 6 (*spatial*) was given as the utterance -- when considered together with looking at and pointing to the box -- had a clear spatial reference.

The scoring by Word List and the scoring by Coding Scheme were conducted by two scorers for each age group, one primary and one secondary scorer blinded to conditions and time of recording. After being trained together on the scoring manual, the primary scorer scored all participating children, whereas the secondary scorer re-scored 20% of the children. Interrater agreement was very high: 46-month-olds: 97.5%; 35-month-olds: 98.3%.

2.4.1.3 Scoring of control questions. It was simply noted whether the child's answer was correct ('yes' or 'no'). Because the scoring of the children's responses to the control questions was straightforward, they were scored online by the experimenter, and rescoring was not considered necessary.

2.5 Results

2.5.1 Preliminary analyses.

Preliminary analyses were conducted in order to ensure that the children across the two event conditions were equivalent with regards to vocabulary. An ANOVA with Age Group (35- vs. 46-month-olds) and Condition (Teddy vs. Game) as between-subjects factors and the children's productive vocabulary as the dependent variable revealed a main effect of Age

Notice: This is the author's version of a work that was accepted for publication in *Consciousness and Cognition.* A definitive version was subsequently published in *Consciousness and Cognition, 55, 91-105.* DOI: 10.1016/j.concog.2017.08.001

Group (F[1, 103] = 36.42, p < .001, η_p^2 = .261). The 46-month-olds had significantly higher productive vocabulary (M_{words} = 652.59, SD = 65.29) than their younger peers (M_{words} = 556.50, SD = 92.58). Importantly, there was no effect of Condition (F = 0.02) and no interaction between Age Group and Condition (F = 0.83), indicating that the children's vocabulary did not differ across conditions.

2.5.2 Outline of analyses

The central part of the data was the scorings from the two-minute waiting periods at T_1-baseline and T_2-test. Three questions were crucial in relation to our hypotheses:

1. Did the number of possible spontaneous verbal utterances obtained at the T_2-test indicating memory of participation in the event differ reliably from the equivalent data obtained at the T_1-baseline?
2. Were the possible spontaneous verbal utterances *congruent* or not with the specific event (Teddy *or* Game), the child was shown at T_1-baseline?
3. Did age affect the results?

We define a *congruent* spontaneous utterance as referring specifically to the event (Teddy *or* Game) that the child saw at T_1-baseline (or by a condition *un*specific utterance just indicating that the child had been there before), and not to the other event, whereas an *incongruent* utterance refers specifically to the event to which the child was *not* exposed at T_1-baseline. Because the analyses focused on congruency, and since preliminary analyses revealed that the two conditions (Teddy or Game) led to very similar results, the analyses were collapsed across conditions (independent samples t-tests revealed that for both the Word List measures as well as for six out of the seven Coding Scheme dimensions, there was no difference between conditions [all ps > .147]; only the *social* dimension showed an effect, t = 4.325, p < .001, r = .38).

Notice: This is the author's version of a work that was accepted for publication in *Consciousness and Cognition.* A definitive version was subsequently published in *Consciousness and Cognition, 55, 91-105.* DOI: 10.1016/j.concog.2017.08.001

Table 1 displays the descriptive statistics based on the congruent spontaneous utterances (both from the Word List and the Coding Scheme) from both events combined, obtained while waiting for the experimenter to return at the T_2-test, as well as the results from the t-tests. Incongruent hits are not displayed, simply because there were none. Note further that the measures obtained at T_1-baseline (congruent and incongruent hits) are not displayed, simply because there were no hits – except for a single incident: This single incident of a word produced at T_1-baseline came from a 35-month-old boy who coincidentally referred to a ball in his mother's bag, and since 'ball' was on the condition specific word list for the Game Conditions this utterance resulted in a random hit.

For both the Word List measures and the Coding Scheme measures, the analyses based on the congruent hits obtained at the T_2-test were analyzed by simple one-sample t-tests tested against the test value of "0" (i.e., "no hits"), because no hits were obtained at T_1-baseline. The only exception was for the single incident mentioned above in the Word List analysis. In order to take this single incident into account for the analysis of the Word List measure, the one-sample t-test for the 35-month-olds was conducted against the exact mean obtained at T_1 in this cell (i.e., slightly above zero). The age groups were contrasted by means of independent samples t-test. All tests were two-tailed. In addition, we analyzed the results from the control questions.

2.5.3 Analyses by Word List.

The children spontaneously produced a highly significant number of condition specific words at the T_2-test indicating memory of participating in the study previously, relative to at T_1-baseline where only one single condition-relevant word was produced by coincidence. No *in*congruent words were uttered spontaneously; thus, none of the spontaneously produced words from the Word List, referred to the event the children had not seen at T_1, indicating

Notice: This is the author's version of a work that was accepted for publication in *Consciousness and Cognition.* A definitive version was subsequently published in *Consciousness and Cognition, 55, 91-105.* DOI: 10.1016/j.concog.2017.08.001

that no false positives were obtained. Methodologically, this means that the attempt to construct two distinct events (the Teddy and the Game event) for which the relevant words to describe them did not overlap, had succeeded.

Importantly, the independent samples t-test with age as grouping variable revealed that the age of the children had no significant effect (see Table 1, Word List row). Although the 46-month-olds tended to produce more congruent words at T_2 relative to their younger peers, the difference was only marginally significant.

Finally, we analyzed whether there was a correlation between the children's productive vocabulary and their spontaneous recall at T_2. The results revealed that there was no relationship regardless of whether we looked at both age groups combined (Pearson r = .10, p = .295 [two-tailed]) or when looking at each age group separately (35-months-olds: Pearson r = .12, p = .372; 46-month-olds: Pearson r = -.13, p=.394, all two-tailed).

2.5.4 Analyses by Coding Scheme.

For the seven dimensions in the Coding Scheme, we used the same strategy of analyses as employed above. Thus, for each of the seven dimensions in the Coding Scheme ([1] *language*, [2] *gestures*, [3] *reliving*, [4] *action details*, [5] *object details*, [6] *spatial details*, and [7] *social details*) we conducted one-sample t-tests tested against the test value of "0" (i.e., "no hits"), because no hits were obtained at T_1-baseline.[2] As can be seen, the pattern of results is highly systematic and closely resembles the results obtained with the Word List measure (see Table 1). For all of the seven dimensions, the analyses revealed that the children in both age groups systematically and reliably produced more hits at T_2 relative to T_1,

[2] Recall, that one 35-month-old child by coincidence mentioned a ball from his mother's bag during the T_1-baseline session, which led to a hit by the Word List measure. Because the scorings by Coding Scheme involved interpretation on the scorers' side (i.e., did the utterance refer to the events?), this incident did not result in a scoring in the Coding Scheme measure, as it was evident that this child was not referring to any of the two events.

Notice: This is the author's version of a work that was accepted for publication in *Consciousness and Cognition.* A definitive version was subsequently published in *Consciousness and Cognition, 55, 91-105.* DOI: 10.1016/j.concog.2017.08.001

indicating that they remembered the event experienced at T_1. Resembling the pattern in results obtained by the Word List measures, no *in*congruent scores were registered on any of the seven dimensions.

Again, independent samples t-tests were used to test for differences between the two age groups. For five (*gesture, action, object, spatial,* and *social*) out of the seven dimensions there were no effects of age group. However, for the remaining two dimensions (*language* and *reliving*) the older children had significantly higher scores (see Table 1 for details). Thus, overall, age group had surprisingly little effect on the results based on the Coding Scheme measures.

2.5.5 Additional Analyses

While we considered the analyses of the means presented in Table 1 to be the core of the present study, we also added some additional information in order to present a more comprehensive picture of the children's spontaneous utterances. First, specific examples of some of the most elaborate spontaneous utterances obtained at the T_2-test are shown in Appendix B.

Second, as is evident from the reported standard deviations (see Table 1), there was large heterogeneity in the children's spontaneous utterances. Thus, in order to elaborate further on the distribution of the frequencies of congruent hits obtained at T_2, Table 2 displays an overview of the number of reported hits from both the Word List measures and the Coding Scheme measures for both age groups.

The numbers reveal that heterogeneity was displayed in both age groups. To examine whether the distribution of hits differed systematically across the two age groups, non-parametric age-wise comparisons were conducted for the eight data points in which the

Notice: This is the author's version of a work that was accepted for publication in *Consciousness and Cognition.* A definitive version was subsequently published in *Consciousness and Cognition, 55, 91-105.* DOI: 10.1016/j.concog.2017.08.001

children had zero hits (one comparison for the Word List measure and one from each of the seven Coding Scheme dimensions). These eight data points were chosen, because they made sense conceptually as they contrasted 'hits' with 'no hits' between the two age groups. By means of Chi-square tests using 2 x 2 contingency tables, significant age differences were obtained in three out of the eight comparisons (1. Word List, $\chi^2[1, N = 110] = 5.54$, $p = .019$; 2. Language dimension, $\chi^2[1, N = 110] = 5.54$, $p = .019$; and 3. Reliving dimension, $\chi^2[1, N = 110] = 5.54$, $p = .019$). The effects for Reliving and Language replicated the parametric results (Table 1), and the effect for Word List was consistent with the trend observed when using parametric analyses (Table 1). Thus, in large, the non-parametric analyses replicated the t-tests, showing only small differences between the age groups.

To summarize, the results from both Word List measures as well as the Coding Scheme measures showed that although there was substantial heterogeneity, the children did indeed talk spontaneously about the event they had been exposed to during the encoding at T_1, whereas they did not talk at all about the event they had not seen (i.e., no false positives). This pattern in results was fairly stable across age. When using the Word List measure, there were no difference between the means of the two age groups (but a trend was observed), and when employing the Coding Scheme measure, there were no difference between the means of the two age groups for five out of the seven dimensions. Thus, in general the 35-month-old children were largely as likely as their 46-month-old peers to spontaneously talk about the event they had seen a week earlier, (and not to talk about the event they had not been exposed to).

2.5.6 Item analysis of Coding Scheme dimensions.

Because the results obtained across the seven Coding Scheme dimensions were highly systematic and clear, and because they addressed memory for the same underlying event, we

Notice: This is the author's version of a work that was accepted for publication in *Consciousness and Cognition.* A definitive version was subsequently published in *Consciousness and Cognition, 55, 91-105.* DOI: 10.1016/j.concog.2017.08.001

expected them to be correlated with one another to some extent and, consequently, that they might be treated as a scale. Reliability analyses were therefore conducted for the seven dimensions (condition-specific and condition unspecific hits combined) on the Congruent scores at the T_2-tests across age groups and events.

The reliability analysis revealed that the seven items (condition-specific and condition unspecific hits combined) had a Cronbach's $\alpha = .917$. Leaving out the item '*object* dimension' would raise the Cronbach's $\alpha = .930$. Leaving out any other items would not raise the Cronbach's α. Thus, we decided to leave out the item '*object* dimension'. Consequently, we ended up with the highly reliable six item scale with a Cronbach's $\alpha = .930$ including the dimensions: *language, gesture, reliving, action, spatial,* and *social.* A total sum score based on these six variables was computed for the subsequent analyses.

First, by means of two one-sample t-tests we analyzed whether the total sum scores based in the six dimensions obtained at T_2 would differ from zero in the two age groups. As expected they clearly did (35-month-olds: $t = 5.49$, $p < .001$, $r = 0.58$; $M_{35\_sum\_score} = 3.43$, $SD = 4.85$; 46-month-olds: $t = 7.24$, $p < .001$, $r = 0.72$, $M_{46\_sum\_score} = 5.00$, $SD = 4.88$).

Second, by means of an independent samples t-test we analyzed whether the two age groups differed with respect to the sum scores. Replicating the pattern in results from the previous analyses, the two age groups did not differ ($t = 1.68$, $p = .09$, $r = 0.16$, two-tailed).

In summary, the item analyses revealed that the Coding Scheme dimensions constituted a highly reliable scale supporting the idea that the dimensions measured the same underlying dimension of episodic remembering. One-sample t-tests based on total sum scores from this six-item scale confirmed and repeated the pattern in results from previous analyses by showing that the children at the T_2-test did indeed spontaneously talk about the event that had seen at T_1-baseline, but did *not* talk about the event they had not seen. Importantly, overall,

Notice: This is the author's version of a work that was accepted for publication in *Consciousness and Cognition.* A definitive version was subsequently published in *Consciousness and Cognition, 55, 91-105.* DOI: 10.1016/j.concog.2017.08.001

age showed no effect, suggesting that the 35-month-old children fared as well as their older peers.

2.5.7 Responses to control questions

By means of a number of non-parametric binomial tests (against p = .50, two-tailed) we analyzed whether the distribution of correct replies regarding the contents of the two boxes across Age Groups as well as the order in which the child was asked about the boxes, was significant. The 'known' box refers to the box to which the child was exposed to the content of during T_1, whereas the 'unknown' box refers to the box to which the child was *not* exposed to the contents of at T_1. The results are displayed in Table 3.

Overall the children responded above chance level when asked whether they knew the contents of the boxes (see Table 3, age groups combined). However, when divided by age it became evident that the 35-month-olds had severe difficulties when asked about the unknown box, that is, the contents of the box they had *not* been shown the contents of at T_1. Regardless of the order of the questions asked, the 35-month-olds never responded above chance level regarding whether they knew the contents of the unknown box (see Table 3).

To analyze this more carefully, we computed a sum score for the number of correct responses regarding whether the children knew the contents of each of the two boxes (the known and the unknown box; range 0-2). By means of an ANOVA with Age Group (35- vs. 46-month-olds) and Order (asked about known box first vs. last) as between-subjects variables and the sum score for the number of correct responses regarding whether the child knew the contents of each the two boxes as the dependent variable, we analyzed whether Age Group and Order affected the number of correct replies. The analysis revealed a main effect of Age Group, $F(1, 103) = 17.33$, $p < .001$, $\eta_p^2 = .144$.

Notice: This is the author's version of a work that was accepted for publication in *Consciousness and Cognition.* A definitive version was subsequently published in *Consciousness and Cognition, 55, 91-105.* DOI: 10.1016/j.concog.2017.08.001

The 46-month-old children obtained significantly higher sum scores ($M_{46\_sum\_score}$ = 1.77, *SD* = 0.43) than their younger peers ($M_{35\_sum\_score}$ = 1.33, *SD* = 0.63). The analyses also revealed a main effect of Order, $F(1, 103) = 5.83$, $p = .018$, $\eta_p^2 = .054$. Across age the children obtained significantly higher sum scores when asked about the known box first (M_{known_first} = 1.64, *SD* = 0.52; M_{known_last} = 1.41, *SD* = 0.63). The analysis revealed no significant interactions.

In summary, the analyses from the control questions revealed that overall the children responded correctly when asked about whether they knew the contents of the two boxes. However, there was a clear difference between the age groups as the 46-month-olds faired significantly better than the 35-month-olds. These results are especially interesting when considering the fact that Age Group generally had considerably less effect on the spontaneous verbal utterances during the T_2-test, as most differences were insignificant.

The contrast in results obtained via the spontaneous recall and the control questions respectively, led us to analyze whether the scores obtained in the two retrieval modes (i.e., spontaneous vs. strategic) correlated. Interestingly, this was not the case, neither for the Word List scores and the sum scores from the control questions (Pearson $r = .02$, $p = .838$ [two-tailed]) nor for the Coding Scheme sum scores and the sum scores for the control questions (Pearson $r = .02$, $p = .875$ [two-tailed]).

Finally, we analyzed whether the sum scores obtained from the control questions correlated with the children's productive vocabulary. Pearson correlations revealed that they did ($r = .22$, $p = .022$, two-tailed).

3.0 Discussion

The results from the present study replicate and expand on previous findings on experimentally induced spontaneous memories (Krøjgaard, et al., 2014). The results were clear, systematic, and in accordance with the proposed hypotheses: When returning to the lab (T_2), the children in both age groups talked spontaneously and reliably about the specific event they had been exposed to at the encoding session (T_1) one week earlier. Spontaneous retrieval was generally not influenced by age, indicating that the 35-month-old children fared as well as the 46-month-olds. The results from the control question were different in this regard. In order to respond to the control questions, the children had to engage in deliberate and strategic recall which requires executive control and involves the use of prefrontal brain areas (e.g., Hall et al., 2014). In contrast to the results obtained from spontaneous recall, the results from the responses to the control questions were affected by the age of the children, and the 46-month-old children reliably outperformed their younger peers. We discuss the theoretical and methodological perspectives related to the results, but before doing so we have to consider (a) whether the memories recorded during the 2 min waiting period at T_2 were actually spontaneous, (b) the distribution of frequencies of the spontaneous memories obtained at T_2, and (c) whether the control questions were adequate.

We defined spontaneous memories as (i) verbally produced, (ii) socially unprompted, and (iii) environmentally cued (Krøjgaard et al., 2014). First, the utterances obviously were verbally produced. Second, the memories produced at T_2 were highly unlikely to be socially prompted, thus, satisfying requirement ii. As previously stated, several precautions had been made in order to ensure that the children were unaware that the study was about memory. None of the children had previously participated in studies in our lab and none of them took part in parallel studies while participating in the present research. Further, when being exposed to their target event at T_1, we had meticulously left out any words or hints that could imply that the study was about memory. Finally, the parents had been carefully informed that it was crucial for the study that the children were unaware of the aim of the study. Thus, we

Notice: This is the author's version of a work that was accepted for publication in *Consciousness and Cognition.* A definitive version was subsequently published in *Consciousness and Cognition, 55, 91-105.* DOI: 10.1016/j.concog.2017.08.001

have no reason to believe that the children's utterances were prompted socially. Third, we have similarly good reasons to assume that the memory utterances were environmentally cued. Although we cannot completely rule out that a few of the children might have been cued internally, that is, by their own thoughts, a number of factors suggest that the children's utterances predominantly were activated by environmental cues. Environmental cues appear to be the dominant sources of spontaneous memories, especially in children (see Krøjgaard et al., 2014, for a review). In addition, besides having an almost complete cue overlap between the to-be-remembered episode and the recall situation, the cues involved were also highly distinct as they possibly referred exclusively to the to-be-remembered event (cf. Watkins & Watkins, 1975). Thus, there are several reasons to assume that the utterances were environmentally cued. In summary, it seems most likely that the memories induced were indeed spontaneous when the children returned to the lab for the second visit.

Whereas the number of the children's spontaneously produced utterances at T_2 were highly significantly different from zero and accompanied by large effect sizes in both age groups, the numerical means appeared somewhat low and with relatively large standard deviations (see Table 1). Thus, in principle the means could reflect quite different distributions across the two age groups, and as such reveal age differences in spontaneous recall. Inspection of the frequencies of hits across measures and age groups showed that the hits were distributed in a heterogeneous manner within both age groups (see Table 2), with the main group difference being that a relatively larger proportion of the 35-month-old children had no spontaneous memories in comparison to their older peers. Additional non-parametric analyses revealed that in three out of eight comparisons, these differences were significant, which largely replicated the parametric analyses.

Notice: This is the author's version of a work that was accepted for publication in *Consciousness and Cognition.* A definitive version was subsequently published in *Consciousness and Cognition, 55, 91-105.* DOI: 10.1016/j.concog.2017.08.001

How should we explain the heterogeneous distribution of hits shown in both age groups? First, it may not be surprising that a relatively large proportion of the children had no spontaneous utterances at all – a tendency that was somewhat more prominent in the youngest age group. Due to the basic design employed, the children were not required to say anything while waiting for the experimenter to return. In addition, the parents had explicitly been instructed not to initiate any conversations, and only to follow up briefly on possible statements from their children. Thus, the conditions under which the study was conducted may have produced a setting in which very little was said. Second, the heterogeneity may, at least partially, reflect a substantial individual variance in the children's inclination to have spontaneous memories. This interpretation is speculative, but it is in accordance with findings with adults showing that there is substantial individual variation in the frequency of spontaneous recollections (e.g., Berntsen, Rubin, & Salgado, 2015; Rasmussen, Ramsgaard, & Berntsen, 2015). Third, some children may simply be more inclined to talk and express their thoughts than others.

Following the design from Krøjgaard et al. (2014), the control questions used in the present study were of a focused yes/no character. One might speculate that the children became confused when asked about the content of the unknown box. This possibility would be consistent with research showing that 3 year old children tend to make more errors when responding to questions for which there is uncertainty regarding the correct answer compared to 4 year olds (e.g., Fernback, Macris, & Sobel, 2012). In addition, 3 year old children often display more difficulties than 4-year-olds when asked about counterfactuals (i.e., "What if...?"; see e.g., Beck, Robinson, Carrol, & Apperly, 2006). However, it should be noted that the control questions employed in the present study did not concern neither uncertainty, nor counterfactuals. Further, the questions were straightforward and did not involve neither

Notice: This is the author's version of a work that was accepted for publication in *Consciousness and Cognition.* A definitive version was subsequently published in *Consciousness and Cognition, 55, 91-105.* DOI: 10.1016/j.concog.2017.08.001

negations (e.g., Dale, Loftus, & Rathbun, 1978) nor misleading phrases (e.g., Cassel, Roebers, & Bjorklund, 1996), which both are known to induce errors in children's responses. Finally, the children in the two age groups combined responded significantly above chance when asked about the content of both boxes, suggesting that the control questions were suitable for the age groups. Thus, we have reason to believe that the obtained age differences are indicative of age related mnemonic difficulties young children have when required to retrieve strategically.

3.1. Theoretical and methodological implications

The results presented here replicate the recent finding that spontaneous memories can be induced experimentally (Krøjgaard, et al., 2014). In addition, the results expand the previously obtained findings in two respects: First, in contrast to Krøjgaard et al. (2014) the to-be-remembered event was a unique one-time occurrence. The event was also induced under substantially more controlled conditions than in the reference study, that is, all children had visited the lab for the same number of times, none of the children knew that the study was about memory, and the event specific props were invisible at test. Thus, it seems safe to assume that the obtained memories were indeed retrieved spontaneously rather than produced by demand characteristics. Second, in the present study, spontaneous memories were induced not only in 46-month-old children, but also in a group of children who, on the average, were below the age of three.

The new and improved version of the design employed in the present study offers at least three advantageous features. First, the present methodology proved highly sensitive for the task (no false positives were obtained). Second, the methodology offers ample opportunities for further in-depth investigations of the determining factors involved in spontaneous retrieval. In the present study, the overlap in distinct cues between the to-be-

Notice: This is the author's version of a work that was accepted for publication in *Consciousness and Cognition.* A definitive version was subsequently published in *Consciousness and Cognition, 55, 91-105.* DOI: 10.1016/j.concog.2017.08.001

remembered episode and the recall situation was almost complete as the experimenter, the spatial location, and the props involved were identical and present at both encounters. However, we do not know whether it was the large number of distinct cues that induced the spontaneous memories, or whether any of these distinct cues were more effective than others. The design allows for future meticulous manipulation of each of the candidate cues in order to examine which cues, or combination of cues, that may be the most potent for inducing spontaneous memories (see Sonne, Kingo, Berntsen, & Krøjgaard, in press). Without the development of the present design or similar, such examinations would be beyond scientific scrutiny.

Third, the development of the dimensions used for the Coding Scheme is a promising new tool for future studies: (a) the dimensions reflect many of the phenomenological qualities conceptualized as 'internal details' in the codings developed by Levine et al. (2002) in order to distinguish between episodic and semantic features of event memories in adults, and (b) six of these dimensions formed a single scale with a high internal consistency, which is advantageous for future research.

At a more general level the present results raise the possibility that the difficulties young children at times display recalling previously experienced episodes (e.g., Dahl et al., 2015; Simcock & Hayne, 2002), may be largely a retrieval problem rather than a memory trace problem, especially at short retention intervals. Hypothetically, young children may at times actually have encoded and consolidated the memory traces of previously experienced episodes, while having severe difficulties retrieving such memories by means of deliberate, strategic recall. This interpretation is in accordance with the contrasting findings in the present study with regard to the effect of age when considering the results from the control questions and the results obtained from spontaneous retrieval, respectively. When asked

Notice: This is the author's version of a work that was accepted for publication in *Consciousness and Cognition.* A definitive version was subsequently published in *Consciousness and Cognition, 55, 91-105.* DOI: 10.1016/j.concog.2017.08.001

directly about the contents of the 'unknown' box, the 35-month-olds (but not the 46-month-olds) probably became confused and did not respond above chance level. However, when talking spontaneously about the contents of the boxes, the 35-month-olds never displayed any confusion regarding which boxes contained what: There were no incidents in which the contents of the box, that the child was exposed to during encoding, were attributed to the 'wrong' box while retrieving the episode spontaneously. This may suggest that the young children's source memory was better during spontaneous than strategic recall. This suggestion is also consistent with the finding that there was no correlation between the results obtained via spontaneous recall and strategic recall respectively, suggesting different underlying mechanisms supporting the two types of recall. Further, whereas the children's productive vocabulary was unrelated to their inclination to have spontaneous recall, productive vocabulary was related to the sum scores obtained when asked control questions requiring strategic retrieval. Thus, the results from the present study support the idea that the development of spontaneous recall of past episodes may precede strategic recall (Berntsen, 2009).

Following the argument above, the present results may challenge the prevailing understanding that episodic memory is a late developmental achievement (Tulving, 2005). From a controlled experiment we have provided evidence that children even below the age of three spontaneously talk about previously experienced episodes. Episodic memories appearing earlier in the ontogenesis than claimed by Tulving (2005) is consistent with previous studies on young children's verbally recalled memories of real-world events (e.g., Fivush, Gray, & Fromhoff, 1987; Hamond & Fivush, 1991) as well as with studies on older children's verbally recalled memories of events experienced in infancy and young childhood in older children (Bauer & Larkina, 2013; Hayne, Gross, McNamee, Fitzgibbon, & Tustin,

Notice: This is the author's version of a work that was accepted for publication in *Consciousness and Cognition.* A definitive version was subsequently published in *Consciousness and Cognition, 55, 91-105.* DOI: 10.1016/j.concog.2017.08.001

2011; Tustin & Hayne, 2010). In addition, Scarf, Gross, Colombo, and Hayne (2013) recently employed the 'spoon test' (i.e., Tulving's [2005] proposal of a non-verbal task to assess episodic memory) to 3- and 4-year-old children. The results revealed that whereas the 4-year-olds passed the test, the 3-year-olds only passed when no delay was inserted.

The results from the present study add to the accumulating and converging evidence that, at least under some circumstances, episodic memories may appear earlier in the ontogenesis than claimed by Tulving (2005). For instance, the 35-month-old girl in example #9 (see Appendix B) spontaneously referred to the specific episode in which she *learned* about the contents of the red box (i.e., "we opened and there were teddies inside"). The girl accompanied her spontaneous verbal utterance by demonstrating with her arms how the doors were opened. In addition, she clapped her hands when referring to the clapping teddies. Thus, this girl not only knew about the contents of the box; she appeared to remember and relive the specific episode in which she learned about it. In our view, the combined and converging verbal and motor behavior by this girl is likely to reflect genuine remembering of the specific episode, or in Tulving's (2005) terms, 'autonoetic awareness'. Note further, that the spontaneous utterances obtained in the present study cannot be explained by recognition; because all the scorings referred to items that were out of view at test, genuine recall was necessary.

As stated above, such memories have been documented in young children below the age of four in several studies using different methodologies. However, in contrast to previous studies, we offer a novel explanation related to the mechanisms of retrieval: By taking departure in the same to-be-remembered material, we show that age has only little effect when retrieval is spontaneous, whereas younger children are worse off than their older peers when using strategic retrieval.

Notice: This is the author's version of a work that was accepted for publication in *Consciousness and Cognition.* A definitive version was subsequently published in *Consciousness and Cognition, 55, 91-105.* DOI: 10.1016/j.concog.2017.08.001

In contrast to Tulving's (2005) claim that episodic memory is a late developmental achievement, phylogenetically as well as ontogenetically, the results from the present study support the recent claim by Berntsen (2009) that spontaneous episodic memories may instead be considered a basic mode of memory that is a developmental precursor for strategic recall (also see Berntsen & Rubin, 2012). This claim is based on the argument that retrieval by means of simple associations -- which seem to be the primary mechanism of retrieval in spontaneous memories, but not in deliberate, strategic retrieval -- are likely to be present (a) early in the ontogenesis, and (b) in other and less cognitively advanced species. In the present study we provide compelling evidence in support of the first part of the claim. Recent evidence from comparative psychology in which apes have been shown to recall episodes across very long retentions intervals when exposed to highly distinctive cues (Martin-Ordas et al., 2013) may be seen as a first step in providing further converging evidence in support of Berntsen's second claim.

A number of limitations should be acknowledged: First, although some of the spontaneously produced utterance were likely to be episodic in Tulving's (2005) sense (cf. Appendix B), many of the utterances were simply too short to allow us to code whether they were truly episodic or not. Second, one might ask to what extent the two measures, the Word List and Coding Scheme, captured the available episodic content of the spontaneous utterances produced. Although the inclination to produce spontaneous memories was unrelated to the children's productive vocabulary in the present study, the children's language abilities (or lack of) may eventually restrict the possibility of capturing spontaneous recall as defined here in case even younger children were examined by means of the present paradigm. Third, we do not know whether the children talked about the events at home between visits. Such information might help to clarify the heterogeneous distribution of hits. Fourth, one

Notice: This is the author's version of a work that was accepted for publication in *Consciousness and Cognition.* A definitive version was subsequently published in *Consciousness and Cognition, 55, 91-105.* DOI: 10.1016/j.concog.2017.08.001

might speculate whether the children would perform better if the control questions had been open-ended. We are currently investigating this possibility. And finally, the demonstration of the events was highly manualized to ensure consistency across trials. We do not know whether manipulating the engagement of the experimenter during the demonstration of the event would affect later recall. Neither do we know whether the children's engagement had effect on the results.

To conclude, we have demonstrated that spontaneous memories of a unique episode can be induced experimentally in 35-month-old children. The results revealed that age had little effect on spontaneous retrieval, whereas age reliably affected the children's ability to respond correctly to the control questions requiring strategic and deliberate recall, as the 46-month-old children outperformed their younger peers on the control questions. Thus, spontaneous retrieval of episodic memories may developmentally precede strategic retrieval. In addition, the present research and its methodology suggest promising avenues for future research on the development of event memory in young children.

Acknowledgement

This research was supported by the Danish National Research Foundation (Grant DNRF89) and Velux Fonden (Grant 10386). We would like to thank Tirill Fjellhaugen, Trine Ravn Andersen, Sidsel Tornhøj Sørensen, and Anne Nørgaard for help with recruiting the participants and scoring the data. We would also like to thank the children who participated in this study and their parents for letting them participate.

Notice: This is the author's version of a work that was accepted for publication in *Consciousness and Cognition.* A definitive version was subsequently published in *Consciousness and Cognition, 55, 91-105.* DOI: 10.1016/j.concog.2017.08.001

References

Ashmead, D.H., & Perlmutter, M. (1980). Infant memory in everyday life. *New Directions for Child and Adolescent Development, 10,* 1-16. DOI: 10.1002/cd.23219801003

Bauer, P., Kroupina, M.G., Schade, J.A., Dropnik, P.L., & Wewerka, S.S. (1998). If memory serves, will language? Later verbal accessibility of early memories. *Development and Psychopathology, 10,* 655-679. DOI: 10.1017/S0954579498001801

Bauer, P.J., & Larkina, M. (2013). The onset of childhood amnesia in childhood: A prospective investigation of the course and determinants of forgetting of early-life events. *Memory, 22*, 907-924. DOI: 10.1080/09658211.2013.854806

Bauer, P.J., van Abbema, D.L., Wiebe, S.A., Cary, M.S., Phill, C., & Burch, M.M. (2004). Props, not pictures, are worth a thousand words: Verbal accessibility of early memories under different conditions of contextual support. *Applied Cognitive Psychology, 18,* 373-392. DOI: 10.1002/acp.1006

Bauer, P.J., Wenner, J.A., & Kroupina, M.G. (2002). *Journal of Cognition and Development, 3*, 21-47. DOI: 10.1207/S15327647JCD0301_3

Bauer, P.J., & Wewerka, S.S. (1995). One- to two-year-olds' recall of events: The more expressed, the more impressed. *Journal of Experimental Child Psychology, 59*, 475-496. DOI:10.1006/jecp.1995.1022

Beck, S.R., Robinson, E.J., Carroll, D.J., & Apperly, I.A. (2006). Children's thinking about counterfactuals and future hypotheticals as possibilities. *Child Development, 77*, 413-426. DOI: 10.1111/j.1467-8624.2006.00879.x

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

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

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

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., & Rubin, D.C. (2012). Understanding autobiographical memory: An ecological theory. In D. Berntsen & D.C. Rubin (Eds.). *Understanding autobiographical memory: Theories and approaches*. Cambridge: Cambridge University Press.

Berntsen, D., Rubin, D.C., & Salgado, S. (2015). The frequency of involuntary autobiographical memories and future thoughts in relation to daydreaming, emotional distress, and age. *Consciousness and Cognition*, *36*, 352-372. DOI: 10.1016/j.concog.2015.07.007

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

Cassel, W.S., Roebers, C.E.M., & Bjorklund, D.F. (1996). Developmental patterns of eyewitness responses to repeated and increasingly suggested questions. *Journal of Experimental Child Psychology*, *61*, 116-133. DOI: 10.1006/jecp.1996.0008

Dahl, J.J., Kingo, O.S., & Krøjgaard, P. (2015). The Magic Shrinking Machine revisited: The presence of props at recall facilitates memory in 3-year-olds. *Developmental Psychology, 51*, 1704-1716. http://dx.doi.org/10.1037/dev0000050

Dale, P.S., Loftus, E.F., & Rathbun, L. (1978). The influence of the form of the question of the eyewitness testimony of preschool children. *Journal of Psycholinguistic Research, 74*, 269-277. DOI: 10.1007/BF01068110

Fernbach, P.M., Macris, D.M., & Sobel, D.M. (2012). Which one made it go? The emergence of diagnostic reasoning in preschoolers. *Cognitive Development, 27,* 39-53. DOI: 10.1016/j.cogdev.2011.10.002

Fivush, R., Gray, J.T., & Fromhoff, F.A. (1987). Two-year-olds talk about the past. *Cognitive Development, 2,* 393-409. http://dx.doi.org/10.1016/S0885-2014(87)80015-1

Hall, S.A., Rubin, D.C., Miles, A., Davis, S.W., Wing, E.A., Cabeza, R., & Berntsen, D. (2014). The neural basis of involuntary episodic memories. *Journal of Cognitive Neuroscience*, *26*, 2385-2399. DOI: 10.1162/jocn_a_00633

Hamond, N.R., & Fivush, R. (1991). Memories of Mickey Mouse: Young children recount their trip to Disneyworld. *Cognitive Development*, *6,* 433-448. http://dx.doi.org/10.1016/0885-2014(91)90048-I

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

Hayne, H., Gross, J., McNamee, S., Fitzgibbon, O., & Tustin, K. (2011). Episodic memory and episodic foresight in 3- and 5-year-old children. *Cognitive Development, 26,* 343-355. http://dx.doi.org/10.1016/j.cogdev.2011.09.006

Hayne, H., Scarf, D., & Imuta, K. (2015). Childhood memories. In J. D. Wright (Ed.), *International encyclopedia of the social & behavioral sciences*. (pp. 465-470). Oxford, UK: Elsevier. http://dx.doi.org/10.1016/B978-0-08-097086-8.51025-3

Hudson, J.A. (1990). The emergence of autobiographical memory in mother-child conversation. In R. Fivush, & J.A. Hudson (Eds.), *Knowing and remembering in young children* (pp. 166-196). New York: Cambridge University Press.

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

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. http://dx.doi.org/10.1016/j.concog.2014. 07.017

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

Martin-Ordas, G., Berntsen, D., & Call, J. (2013). Memory for distant past events in chimpanzees and orangutans. *Current Biology*, *23,* 1438-1441. http://dx.doi.org/10.1016/j.cub.2013.06.017

Nelson, K. (1989). *Narratives from the crib.* Cambridge MA: Harvard University Press.

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. DOI: 10.1002/cd.23219801008

Newcombe, N.S., Lloyd, M.E., & Ratliff, K.R. (2007). Development of episodic and autobiographical memory: A cognitive neuroscience perspective. In R.V. Kail (Ed.), *Advances in child development and behavior* (Vol. 35, pp. 37-85). San Diego, CA: Elsevier.

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

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

Scarf, D., Gross, J., Colombo, M., & Hayne, H. (2013). To have and to hold: Episodic memory on 3- and 4-year old children. *Developmental Psychobiology, 55*, 125-132. DOI: 10.1002/dev.21004

Schlagman, S., & Kvavilashvili, L. (2008). Involuntary autobiographical memories in anoutside the laboratory: how different are they from voluntary autobiographical memories? *Memory & Cognition, 36*, 920-932. DOI: 10.3758/MC.36.5.920

Simcock, G., & Hayne, H. (2002). Breaking the barrier? Children fail to translate their preverbal memories into language. *Psychological Science, 13,* 225-231. DOI: 10.1111/1467-9280.00442

Sonne, T., Kingo, O.S., Berntsen, D., & Krøjgaard, P. (in press). Thirty-five-month-old children have spontaneous memories despite change of context for retrieval. *Memory.*

Todd, C.M., & Perlmutter, M. (1980). Reality recalled by preschool children. *New Directions for Child and Adolescent Development*, *10,* 69-85. DOI: 10.1002/cd.23219801007

Tulving, E. (2005). Episodic memory and autonoesis: Uniquely human? In H.S. Terrace and J. Metcalfe (Eds.), *The missing link in cognition: Origins of self-reflective consciousness* (pp. 3–56). Oxford: Oxford University Press.

Tustin, K., & Hayne, H. (2010). Defining the boundary: Age-related changes in childhood amnesia. *Developmental Psychology, 46,* 1049-1061. DOI: 10.1037/a0020105

Watkins, O.C., & Watkins, M.J. (1975). Buildup of proactive inhibition as a cue-overload effect. *Journal of Experimental Psychology: Human Learning and Memory*, *104*, 442-452. http://psycnet.apa.org/doi/10.1037/0278-7393.1.4.442

Notice: This is the author's version of a work that was accepted for publication in *Consciousness and Cognition.* A definitive version was subsequently published in *Consciousness and Cognition, 55, 91-105.* DOI: 10.1016/j.concog.2017.08.001

Appendix A

Word List in alphabetical order for each of the two conditions as well as a 'Unspecific' list of words each indicating that the child had been there before without being specific about any of the two conditions. Note that synonyms and conjugations of the words presented in the lists counted as well.

Teddy Condition	**Game Condition**	**Unspecific**
Alfred	Ball	'Always'
Animal	Blue	'Have been'
Button	Bowling	'Again'
Clap	Bucket	'Last time'
Dog	Game	
Elephant	Green	
Elly	Hit	
Flip	Medal	
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 *Consciousness and Cognition.* A definitive version was subsequently published in *Consciousness and Cognition, 55, 91-105.* DOI: 10.1016/j.concog.2017.08.001

Appendix B.

Presents examples of spontaneous utterances obtained at the T_2-test.

#	*Age Grp*	*Exact age*	*Gender*	*Condition*	*Quote*
1	46	48.86	boy	Teddy	"There are animals inside [pointing at the red box]. I saw that the other day I was here. Then I saw, there were animals. Two. They were not alive at all."
2		46.43	girl	Teddy	"Why does she [referring to the experimenter] always have to leave? She has to sing those teddies [pointing at the red box]."
3		46.53	boy	Teddy	"Now she has done that twice [just after the experimenter left the room again]. There are animals inside [pointing at the red box]. She will also open that box today. Mom, the elephant is inside. There is the teddy. There is the dog […]. We have been here twice now."
4		46.36	girl	Game	"What will she [referring to the experimenter] do today? She has other clothes on today. It's nicer than the other [presumably referring to the clothes the experimenter were wearing at the first visit]. […] Once, I got two medals."
5		49.16	girl	Game	"And then I tried that game [pointing at the brown box], there was also a Bowling game. And then she [referring to the experimenter] asked me, if I could say the words she said."
6		47.30	boy	Teddy	"I've been here before. We will try the funny one again – the funny ones, that lives inside [pointing at the red box]. Last time I was here, I thought it was a mail box* [pointing at the red box]."
7		46.46	boy	Game	"There was a really funny game. There was a ball and one where I should turn over that one. But I couldn't throw all those balls into the buckets."
8	35	35.80	boy	Teddy	[Walks to the red box and refers to it] "I want to turn it on today" [presumably referring to the mechanical teddies than can be turned on]
9		35.77	girl	Teddy	"There's two teddies inside [pointing at the red box], but we opened [demonstrating with her arms how to open the doors of the red box] and there were teddies inside. They clapped [she claps herself while speaking]
10		35.50	girl	Teddy	"Last time we did here, the lady came again. That was bad."
11		35.20	boy	Teddy	"When the girl comes, she will bring a key and open [to] the two teddies."

Notice: This is the author’s version of a work that was accepted for publication in *Consciousness and Cognition.* A definitive version was subsequently published in *Consciousness and Cognition, 55, 91-105.* DOI: 10.1016/j.concog.2017.08.001

12	35.23	boy	Teddy	“There is an elephant inside. There is no key. She [referring to the experimenter] will come and open in a minute.”

*I Denmark mail boxes are red and made of metal.

Notice: This is the author's version of a work that was accepted for publication in *Consciousness and Cognition.* A definitive version was subsequently published in *Consciousness and Cognition, 55, 91-105.* DOI: 10.1016/j.concog.2017.08.001

Table 1.

Descriptive statistics and t-tests based on the Congruent Word List measures and the Congruent Coding Scheme measures at T_2 for both events combined across Age Groups.

	35-month-olds (n= 60)							46-month-olds (n= 50)							35-month-olds vs. 46-month-olds			
	Descriptives		One-sample t-tests (tested against the value *x*)					Descriptives		One-sample t-tests (tested against the value *x*)					Independent samples t-test			
Measure	*M*	*SD*	*x*	*t*	*df*	*p*	*r*	*M*	*SD*	*x*	*t*	*df*	*p*	*r*	*t*	*df*	*p*	*r*
Word List	1.00	1.43	.02	5.32	59	<.001	0.57	1.56	1.68	0	6.57	49	<.001	0.68	-1.89	108	= .061	0.18
Coding Scheme																		
Language	0.82	1.16	0	5.47	59	<.001	0.58	1.32	1.30	0	7.18	49	<.001	0.72	-2.15	108	= .034	0.20
Gesture	0.45	0.72	0	4.82	59	<.001	0.53	0.64	0.78	0	5.83	49	<.001	0.64	-1.33	108	= .187	0.13
Reliving	0.82	1.16	0	5.47	59	<.001	0.58	1.28	1.28	0	7.08	49	<.001	0.71	-1.99	108	= .049	0.19
Action	0.38	0.69	0	4.30	59	< .001	0.49	0.66	0.87	0	5.36	49	< .001	0.61	-1.86	108	= .066	0.18
Object	0.13	0.34	0	3.01	59	= .004	0.37	0.26	0.57	0	3.27	49	= .002	0.42	-1.49	108	= .151	0.14
Spatial	0.57	0.83	0	5.28	59	<.001	0.57	0.62	0.70	0	6.29	49	<.001	0.67	-0.36	108	= .719	0.03
Social	0.40	0.74	0	4.18	59	<.001	0.48	0.48	0.79	0	4.30	49	<.001	0.52	-0.55	108	= .585	0.05

Note: The Congruent hits are based on the condition specific hits (e.g., "teddy", "elephant") as well as condition unspecific hits (e.g., "last time", "again"). All t-test are two-tailed. Note further that the Word List variable for the 35-month-olds was not tested against "0" in the One-sample t-tests. The reason is that for this variable, there was a single false positive score during baseline originating from the one boy who coincidentally mentioned a ball from his mother's bag.

Notice: This is the author's version of a work that was accepted for publication in *Consciousness and Cognition.* A definitive version was subsequently published in *Consciousness and Cognition, 55, 91-105.* DOI: 10.1016/j.concog.2017.08.001

Table 2.
Displays the frequencies of Congruent hits at T_2 from the Word List as well as the Coding Scheme measures in both Age Groups.

	35-month-olds								46-month-olds							
	Frequencies of hits on the Words List as well as the Coding Scheme*								Frequencies of hits on the Words List as well as the Coding Scheme*							
Number of hits	0	1	2	3	4	5	6	Total	0	1	2	3	4	5	6	Total
Word List	35	7	8	4	5	1	0	60	17	14	6	6	2	4	1	50
Coding Scheme																
Language	35	10	8	5	2	0	0	60	17	15	7	7	4	0	0	50
Gesture	40	14	5	1	0	0	0	60	26	17	6	1	0	0	0	50
Reliving	35	10	8	5	2	0	0	60	17	16	7	6	4	0	0	50
Action	44	9	7	0	0	0	0	60	28	13	7	2	0	0	0	50
Object	52	8	0	0	0	0	0	60	40	7	3	0	0	0	0	50
Spatial	36	16	7	0	4	0	0	60	25	19	6	0	0	0	0	50
Social	44	9	6	1	0	0	0	60	34	9	6	1	0	0	0	50

*****Both condition specific and condition unspecific hits are counted

Notice: This is the author's version of a work that was accepted for publication in *Consciousness and Cognition.* A definitive version was subsequently published in *Consciousness and Cognition, 55, 91-105.* DOI: 10.1016/j.concog.2017.08.001

Table 3.
Descriptives and binominal statistics from the results from for the control questions asked (did the child know the contents of the boxes?) at T_2 across Age Groups.

Age	Box asked about first?	N^{α} Know/ Unknown	Correct replies regarding known box: Count	%	p*	Correct replies regarding unknown box: Count	%	p*
Comb.	Known	53/54	49	92	<.001	39	72	=.001
	Unknown	55/54	44	80	<.001	32	59	=.22
	(Both boxes)	108/108	94	86	<.001	71	66	=.001
35	Known	30/30	26	87	<.001	16	53	=.856
	Unknown	30/30	23	77	=.005	15	50	=1
	(Both boxes)	60/60	49	82	<.001	31	52	=.897
46	Known	23/24	23	100	<.001	23	96	<.001
	Unknown	25/24	21	84	=.001	17	71	=.064
	(Both boxes)	48/48	45	92	<.001	40	83	<.001

α In a few cases we failed to make the children respond to the control questions. Thus, for each test we have specified the total number of obtained responses on which the analyses are were based.
***** Binomial test, test-value: p = .50 (two-tailed).

Notice: This is the author's version of a work that was accepted for publication in *Consciousness and Cognition.* A definitive version was subsequently published in *Consciousness and Cognition, 55, 91-105.* DOI: 10.1016/j.concog.2017.08.001

Figure Captions

Figure 1. What the children could see while waiting.

Figure 2. Schematic representation of the design employed.

Notice: This is the author's version of a work that was accepted for publication in *Consciousness and Cognition.* A definitive version was subsequently published in *Consciousness and Cognition, 55, 91-105.* DOI: 10.1016/j.concog.2017.08.001

Figure 1. What the children could see while waiting.

Notice: This is the author's version of a work that was accepted for publication in *Consciousness and Cognition.* A definitive version was subsequently published in *Consciousness and Cognition, 55, 91-105.* DOI: 10.1016/j.concog.2017.08.001

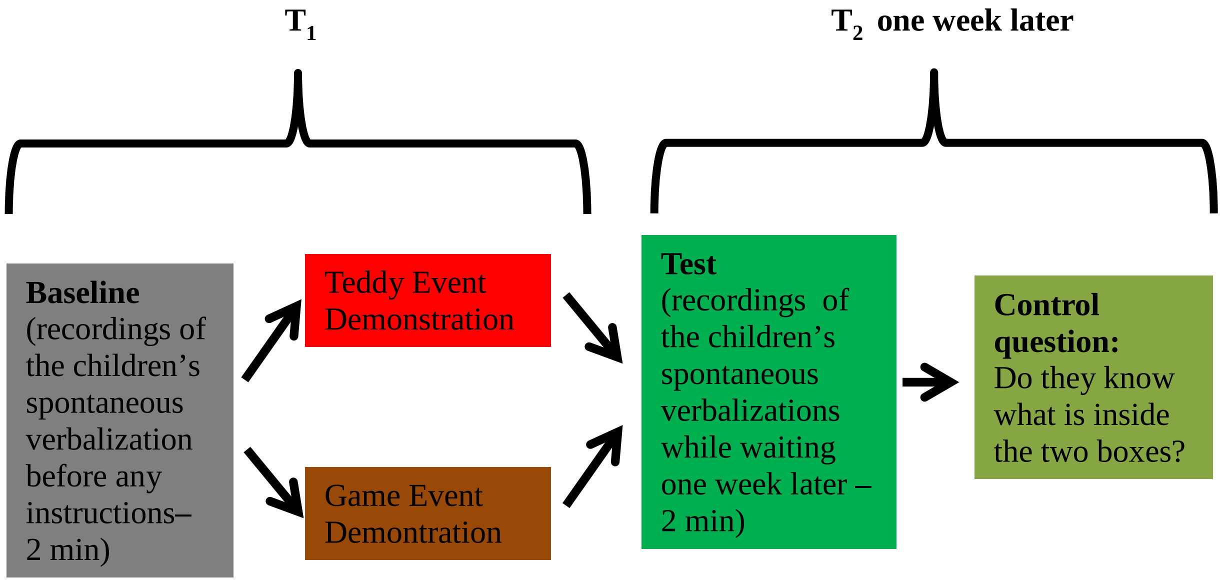

Figure 2. Schematic representation of the design employed.