Notice: This is the author's version of a work that was accepted for publication in *Cognitive Development.* A definitive version was subsequently published in *Cognitive Development, 39*, 168-180. DOI: 10.1016/j.cogdev.2016.06.001

Twenty-four-month-olds' nonverbal memory for expected and unexpected versions of familiar events

Jonna J. Dahl*

Osman S. Kingo†

Peter Krøjgaard†

* *Pedagogical Psychological Counseling, Municipality of Hedensted*

† *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:

Jonna J. Dahl, Pedagogical Psychological Counseling, Municipality of Hedensted, Juelsminde Rådhus, Tofteskovvej 4, DK-7130 Juelsminde, Denmark, Mail: jonna.j.dahl@hedensted.dk, Phone: 0045 30328176.

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

Notice: This is the author's version of a work that was accepted for publication in *Cognitive Development*. A definitive version was subsequently published in *Cognitive Development*, *39*, 168-180. DOI: 10.1016/j.cogdev.2016.06.001

Abstract

Very little is known regarding whether scripted knowledge affects memory in infancy. By means of the elicited imitation paradigm we examined whether 24-month-olds' (N=112) differentially re-enacted expected and unexpected 4th steps of two highly familiar and of two less familiar 5-step events immediately as well as across a two-week retention interval. The results revealed that overall the infants re-enacted fewer unexpected 4th steps (relative to expected 4th steps) of the events as they had been demonstrated to the infants, and that the infants re-enacted fewer 4th steps (expected and unexpected) at the delayed recall test as they had been demonstrated. However, although the degree of familiarity of the events influenced the total number of actions (re)-enacted, familiarity did not affect how often the infants re-enacted the crucial 4th steps as they had been demonstrated. The results are discussed in relation to the prevailing theories of scripted knowledge in infancy.

Notice: This is the author's version of a work that was accepted for publication in *Cognitive Development.* A definitive version was subsequently published in *Cognitive Development, 39,* 168-180. DOI: 10.1016/j.cogdev.2016.06.001

Twenty-four-month-olds' nonverbal memory for expected and unexpected versions of familiar events

1.0 Introduction

The term scripted knowledge or simply *scripts* was introduced by Shank and Abelson (1977). Scripts refer to abstract knowledge structures reflecting the understanding of the temporal and causal sequences of familiar action sequences. As adults we have scripts for numerous familiar events typically experienced in everyday life, such as visiting a restaurant, getting ready to go to work, or celebrating a child's birthday. Scripts help us to free cognitive capacity for other tasks, to interpret experiences, and to predict what may or may not happen in similar future events (Dahl, Sonne, Kingo, & Krøjgaard, 2013).

From a developmental perspective where scripts at times are referred to as general event representations (GERs, Hudson & Mayhew, 2009) we are interested in when and how scripts are learned, and how they may affect memory. In the present study we focused on the latter aspect, that is, how scripted knowledge may affect memory. More specifically, we examined how 24-month-olds re-enacted expected and unexpected incidents of highly familiar and less familiar events by means of the elicited imitation paradigm employing an immediate as well as a (two-week) delayed test. We define *unexpected* incidents as incidents that unambiguously violate the typical course of a particular event. We distinguish between highly familiar and less familiar events by means of the infants' baseline performance: Without motor instruction infants should produce reliably more target actions for the highly familiar events relative to the less familiar events. When considering infants we use the term 'events', because before having actually tested the infants we do not know whether the events qualify as genuine scripts (i.e., abstract knowledge structures reflecting the understanding of the temporal and causal sequences of familiar action sequences) – whereas they almost certainly would for adults.

Young children's recall of expected and unexpected versions of scripted events has been investigated in the verbal domain, while infants' nonverbal re-enactments of highly familiar and less familiar events have been examined by means of the elicited imitation paradigm. To the best of our knowledge, no one has previously examined memory of unexpected versions of highly familiar and less familiar events in nonverbal infants. In the

Notice: This is the author's version of a work that was accepted for publication in *Cognitive Development.* A definitive version was subsequently published in *Cognitive Development, 39,* 168-180. DOI: 10.1016/j.cogdev.2016.06.001

following we briefly outline the two lines of research we set out to combine, that is, verbally proficient children's recall of scripted events, and infants' recall of familiar events.

1.1 Verbally proficient children's recall of scripted events

Groundbreaking research conducted by Katherine Nelson and her colleagues revealed that even three-year-olds possessed stable, ordered event representations of familiar events when asked to verbally report "what happens" in typically encountered events such as going to McDonalds, having lunch at daycare, or having dinner at home (e.g., Nelson 1978, 1979; Nelson & Gruendel, 1981, 1986). These findings were in sharp contrast to the then prevailing view that young children's event memories were unorganized and idiosyncratic (e.g., Piaget & Inhelder, 1973). Although the amount of information reported when asked about "what happens" in a typical event increases as the child grows older, several studies have shown that even young children tend to report the events in a generalized and temporally structured manner (Fivush, 1984; Nelson & Gruendel, 1981, 1986; Hudson & Nelson, 1986). A robust finding is that when young children have multiple experiences with similar events, their event representations become generalized and they have difficulty remembering details of a specific episode of a recurrent event (Hudson, 1990; Hudson, Fivush, & Kuebli, 1992; Kuebli & Fivush, 1994). For example, Hudson and Nelson (1983) examined preschoolers' and first graders' use of scripts in story recall and found that even though first graders recalled unexpected or discrepant information to a greater extent than did preschoolers, both age groups often falsely reported scripted story information indicating that their recall was influenced by their scripted knowledge.

These findings are consistent with the assumption provided by both the event-schema theory (Hudson et al., 1992) and the fuzzy-trace theory (Brainerd & Reyna, 2001; Brainerd, Reyna, Wright, & Mojardin, 2003) suggesting that as a function of repeated experiences young children's memory tends to improve for what usually happens in an event, whereas their recall of specific (including unusual) details decreases (e.g., Hudson & Nelson 1983). According to the fuzzy-trace theory both children and adults develop two fundamentally different types of memory representations: a gist trace which is fuzzy and imprecise, yet represents information concerning common features from across the experienced scripted events, and a verbatim trace representing the factual details associated with the event (Brainerd & Reyna, 2001; Brainerd, Reyna, & Kneer, 1995). The assumption is that as time passes, the gist traces are easier to access compared to the verbatim traces which are more

Notice: This is the author's version of a work that was accepted for publication in *Cognitive Development*. A definitive version was subsequently published in *Cognitive Development*, *39*, 168-180. DOI: 10.1016/j.cogdev.2016.06.001

fragile and rapidly forgotten (Brainerd & Reyna, 2001; Brainerd et al., 2003). Note that this account may explain why young children across delays tend to recall a 'single' episode of familiar events according to the script and not necessarily how the specific episode occurred.

Meanwhile, some studies have revealed that when a particular episode deviates substantially from what usually happens in a familiar event, young children (like older children and adults) seem to remember the atypical information better than information which is in accordance with their script knowledge (Davidson & Hoe, 1993; Hudson, 1990). Furthermore, several studies with preschoolers (e.g., Fivush, Gray, & Fromhoff, 1988; Nelson, 1988) have shown that novel events occurring only once tend to remain relatively distinct in memory. These findings are in accordance with the body of research demonstrating the role of *distinctiveness* on memory (e.g., Hunt & Worthen, 2006). According to this literature, a distinctive event is well remembered because it is surprising, salient, bizarre, or novel, attracting considerable attention leading to enhanced memory (Hunt, 2006). Thus, a distinctive event tends to resemble a given category of events while at the same time violating what usually happens in this given event category – in short: a distinctive event is often considered to be an unexpected exemplar of a recurrent event. However, as emphasized by Howe (2006) even though the role of distinctiveness in memory is well documented in adults and older children, we do not know whether this effect can also be translated to memory in young children and infants. In fact, prior research on how infants recall unexpected incidents of familiar events is almost absent. In the following we review this literature.

1.2 Infants' re-enactment of familiar events

Using the elicited imitation paradigm, a number of studies, most notably by Bauer and her colleagues, have demonstrated ordered recall of both familiar and novel events by 11- to 36-month-olds (Bauer & Mandler, 1989, 1992; Bauer & Shore, 1987; Bauer & Thal, 1990; Bauer & Travis, 1993; O'Connell & Gerard, 1985). In these studies the familiar events were typically chosen among events that one would consider highly salient and frequently encountered in the lives of infants (e.g., being put to bed, having a bath, or having breakfast).

Although, familiarity tended to facilitate ordered recall, infants also recalled novel events after only a single experience (Bauer & Mandler, 1989; Bauer & Thal, 1990). Some studies have revealed that changing the expected order of the event components disturbed the infants' recall (e.g. Fivush & Mandler, 1985; O'Connell & Gerard, 1985). For example,

Notice: This is the author's version of a work that was accepted for publication in *Cognitive Development.* A definitive version was subsequently published in *Cognitive Development, 39,* 168-180. DOI: 10.1016/j.cogdev.2016.06.001

O'Connell and Gerard (1985) tested 20-, 24-, 28-, and 36-month-olds' recall of three types of events: canonical, reversed, and scrambled order and found that 24-month-olds only re-enacted the modeled events when they were presented in their typical order.

Other studies have examined infants' immediate recall of novel events in which irrelevant or unimportant event components were embedded in the to-be-remembered sequence (Bauer, 1992, Experiment 2; Bauer & Mandler, 1989, Experiment 2). Bauer and Mandler (1989, Experiment 2) found that 19-, 25- and 31-month-olds tended to displace or omit an irrelevant event component (i.e., attaching a sticker), particularly when they re-enacted the specific order of the enabling events (i.e., events for which the end goal can only be achieved by conducting the actions in one specific order) compared to when they re-enacted the arbitrarily ordered events (i.e., events for which there are no constrains of the order of the actions involved). However, the two experiments focused on the infants' immediate recall of enabling and arbitrarily ordered novel events and thus did not look at whether the irrelevant components would interfere with their recall across time.

In an additional experiment, Bauer (1992, Experiment 3) investigated, across a two-week delay, 25-month-olds recall of familiar or novel events in which an interrupting element was an integral aspect of an enabling event (i.e., a 'wiping the table' action was inserted in between [a] opening and [b] subsequently throwing away some trash in a trash can). The results revealed that the infants tended to bring together the two split elements in their recall while displacing or omitting the interrupting element. However, when considering the infants' recall of the familiar events, Bauer (1992, Experiment 3) did not investigate whether *atypical* or *unexpected* event components would interfere with recall of highly familiar events.

In some of the studies cited above, baseline measures were obtained for the familiar events (Bauer & Mandler, 1989, 1992; Bauer & Shore, 1987; Bauer & Travis, 1993), whereas in other studies they were not (Bauer & Thal, 1990; O'Connell & Gerard, 1985). The baseline performance of the familiar events in the above mentioned studies did typically not differ from the baseline performance of the novel events ((Bauer & Mandler, 1989, 1992; Bauer & Shore, 1987). Whether infants are actually familiar with a presumably 'familiar' event has previously more or less been taken for granted in the literature. Thus, in most of these studies 'familiarity' has been assumed although not explicitly tested. However, in one study in particular, the researchers meticulously ensured that the familiar events were indeed familiar to the 24-month-old infants: Bauer and Travis (1993, exp. 2) had parents specify

Notice: This is the author's version of a work that was accepted for publication in *Cognitive Development.* A definitive version was subsequently published in *Cognitive Development, 39,* 168-180. DOI: 10.1016/j.cogdev.2016.06.001

distinct events that they were certain that their 24-month-old infants would be familiar with. Although the baseline measures for these familiar events were not explicitly contrasted with novel events as the data from familiar and novel events were collected in two different experiments (exp 1: novel; exp. 2: familiar), the baseline measures of the mean number of actions enacted from the 5-step familiar events were on average numerically higher (2.00 [enabling], and 2.17 [arbitrary]) than the equivalent means from the novel events (0.50 [enabling], and 2.04 [arbitrary]). As enabling and arbitrary relations were controlled across experiments, we take the results from the baseline measures to indicate that the infants probably knew more about familiar events before entering the lab, than they knew about the novel events. When returning to re-enact the demonstrations after a 2-3 day retention interval, the children in the Bauer and Travis (1993) study on average performed *fewer* target actions from the familiar events (3.38 [enabling], and 3.17 [arbitrary]) relative to the novel events (4.42 [enabling], and 4.88 [arbitrary]). However, these results were obtained in two different experiments and the numbers were not contrasted explicitly.

1.3 The present study

Whereas memory for expected and unexpected versions of highly familiar and less familiar events has been investigated in verbally proficient children, these questions have so far not been addressed in infants. In addition, and as already stated, very little is known regarding whether scripted knowledge affect memory across time in infancy. In the present study we attempted to pursue these questions in infancy. As previously stated, we define unexpected incidents as incidents that unambiguously violate the typical course of a particular event. Thus, in order to qualify as an unexpected incident, the script shall not only be temporarily interrupted or have an irrelevant component added (cf. Bauer 1992; Bauer & Mandler, 1989). Instead, one or more of the core components in a scripted sequence needs to be violated.

Unexpected incidents are interesting when considering the possible effect of scripted knowledge on memory, because they constitute a strong test for the possible effect of scripted knowledge on the infants' memory for specific events due to their radical violation of the typical event. In addition, such unexpected incidents are in accordance with what have been coined distinctive events which, as previously mentioned, have not yet been examined in infancy (Howe, 2006). We reasoned that if events highly familiar to infants had already been scripted, then it would be interesting to examine whether these scripts over time would

Notice: This is the author's version of a work that was accepted for publication in *Cognitive Development.* A definitive version was subsequently published in *Cognitive Development, 39,* 168-180. DOI: 10.1016/j.cogdev.2016.06.001

affect the infants' memory of the unexpected incidents. Inspired by prior research (e.g., Bauer & Thal, 1990; Bauer & Travis, 1993; O'Connell & Gerard, 1985) we assumed that 24-month-olds would be highly familiar with events such as "going-to-bed" and "having-a-bath" which the infants encounter on a daily basis, and less familiar with events such as "changing-a-tire-on-a-car" and "going-to-the-Zoo" (Fivush, Gray, & Fromhoff, 1987; Hudson & Nelson, 1986) which probably are less frequently experienced in the lives of infants.

The baseline assessments of each event, in which the number of target actions enacted *prior* to the demonstrations were recorded for each event, allowed us to assess whether the two highly familiar events were indeed more familiar to the infants than the two less familiar events. Note, that in contrast to previous studies using elicited imitation to assess familiarity of events (e.g., Bauer & Travis, 1993) we used explicit, verbal prompts at baseline to encourage the infants to show the experimenter how to enact a certain event (see the procedure section for details). Thus, the infants' explicitly instructed baseline performance was used as a proxy for their scripted knowledge of the events prior to any demonstration provided by the experimenter.

By means of the elicited imitation paradigm we investigated whether 24-month-olds re-enacted presumably familiar events while manipulating three factors experimentally: (1) Familiarity (highly familiar vs. less familiar events), (2) Expectedness (expected vs. unexpected version of the events), and (3) Time of Test (immediate test vs. delayed test [two weeks]). 'Expectedness' was operationalized as whether the 4^{th} step of the 5-step sequences either followed (expected) or violated (unexpected) the script. In order to examine the possible effect of scripted knowledge in particular, we carefully instructed the infants to re-enact the sequences *as they had been demonstrated to them* by the experimenter – regardless of whether the 4^{th} steps were expected or unexpected. Because this was the first study to examine infants' memory of unexpected versions of highly familiar and less familiar events and since the predictions derived from the two prominent theoretical approaches (fuzzy-trace theory vs. distinctiveness theory) differed, we had two contradictory hypotheses: (1) based on the fuzzy-trace theory we would *not* expect the infants to re-enact the unexpected versions of the familiar events over the two-week retention interval. However, (2) based on the distinctiveness theory, we *would* expect the infants to re-enact the unexpected versions of the familiar events over the two-week retention interval.

Notice: This is the author's version of a work that was accepted for publication in *Cognitive Development*. A definitive version was subsequently published in *Cognitive Development, 39*, 168-180. DOI: 10.1016/j.cogdev.2016.06.001

2.0 Method

Using the elicited imitation paradigm, we employed a within-subjects design in which all infants were tested on four different events, each involving five steps. Each event was tested in three phases: *baseline* (without preceding demonstration), *immediate imitation* (immediately after demonstration), and *delayed recall* (after two-week retention). Each event came in two versions: An *expected* version (following the script) in which the 4th step was typical, and an *unexpected* version (violating the script) in which the 4th step was highly unusual (see below).

2.1 Participants

A total of 112 infants (60 females), 24 months of age (M = 24.01, SD = .28, range = 23.60 – 24.50), participated in the present study. Only infants who completed the test for all four events were included in the final sample. An additional 14 infants were tested but excluded from later analysis due to fussiness (seven), experimental error (six) and missing data (one). The children were predominantly Scandinavian Caucasian, living in families with middle to higher SES. All infants were recruited from the Aarhus area in Denmark via registers from the National Board of Health and they were all full term and healthy. Besides receiving a small gift for their participation, no compensation was offered.

2.2 Materials

All infants were seen individually in the same 16 m^2 laboratory room twice, and they were seated in a high chair at a table (180 cm wide and 90 cm deep) in front of the experimenter.

2.2.1 The four events

The four events were: (1) the Sleep Event, (2) the Bath Event, (3) the Car Event, and (4) the Zoo Event. Two of these events (the Sleep Event and the Bath Event) were considered to be *highly familiar* to the infants, whereas the two others (the Car Event and the Zoo Event) were considered to be *less familiar* to the infants (see Fig. 1).

It was invariably the fourth step, and only the fourth step, that came in both an expected and an unexpected version. See Fig. 2 for an example of the expected and unexpected course of actions, here illustrated for the Sleep Event.

Notice: This is the author's version of a work that was accepted for publication in *Cognitive Development.* A definitive version was subsequently published in *Cognitive Development, 39,* 168-180. DOI: 10.1016/j.cogdev.2016.06.001

The four events were modeled and narrated as follows:

The Sleep Event: 1) "I get the toothbrush from the cup", 2) "I put toothpaste on the toothbrush", 3) "I brush the baby's teeth", 4) "Then I put the baby *in* the bed" (Expected version) / "Then I put the baby *under* the bed" (Unexpected version), 5) "And then I put on the duvet".

The Bath Event: 1) "I take of the teddy's clothes", 2) "I put the teddy in the bathtub", 3) "I wash teddy with a sponge", 4) "Teddy comes up to be *dried with a towel"* (Expected version) / "Teddy comes up and *I throw the towel* into the bathtub" (Unexpected version), 5) "Then I put the clothes back on teddy".

The Car Event: 1) "The man gets out of the car", 2) "I lift the car", 3) "I take off the red tire", 4) "I *put on the new tire"* (Expected version) / "I put the new tire *in front of the car"* (Unexpected version), 5) "I lower the car down again".

The Zoo Event: 1) "The boy enters the Zoo", 2) "He sees a giraffe eating leafs from a tree", 3) "He sees an elephant in a cage", 4) "The zoo-keeper *feeds the penguin* with a fish" (Expected version) / "The zoo-keeper *feeds the boy* with a fish" (Unexpected version), 5) "The boy leaves the Zoo, goodbye".

2.3 Procedure

All infants were seen in the laboratory twice separated by two weeks (± 2 days, *M* = 14.07, *SD* = .94), and one female experimenter (the first author) tested all of the infants. The parent was sitting next to the infant and was instructed not to interfere during the test.

Each infant was tested on the four, five-step, events described above. All infants saw all four events: The two highly familiar events (one event in the expected version, and one event in the unexpected version) and the two less familiar events (one event in the expected version, and one event in the unexpected version; see above). The order in which the events were presented as well as whether the version was expected or unexpected was counterbalanced across infants. For the sake of simplicity, we randomized the infants into one of the following four fixed orders (28 infants in each): (1) Bath Event, Expected version – Car Event, Unexpected version – Zoo Event, Expected version – Sleep Event, Unexpected version, (2) Zoo Event, Unexpected version – Bath Event, Unexpected version – Sleep Event, Expected version – Car Event, Expected version, (3) Sleep Event, Unexpected version – Zoo Event, Expected version – Car Event, Unexpected version – Bath Event, Expected version, and (4) Car Event, Expected version – Sleep Event, Expected – Bath

Notice: This is the author's version of a work that was accepted for publication in *Cognitive Development*. A definitive version was subsequently published in *Cognitive Development, 39*, 168-180. DOI: 10.1016/j.cogdev.2016.06.001

Event, Unexpected – Zoo Event, Unexpected. For each infant the order of events was kept the same at both visits.

At the first visit the infants went through the three test phases *baseline, demonstration*, and *immediate imitation*. The baseline served the purpose of providing a measure of the infants' script knowledge prior to the demonstration of each of the four events as well as potential on-line reasoning. At the second visit each infant was subjected to the *delayed recall* test (see below). At both visits all test phases were conducted consecutively for each event before proceeding to the next event. The two sessions were recorded for later off-line scoring (see below) using two JVC Everio digital camcorders. Before beginning the test phases at the first visit, two warm-up events, designed to acquaint the infants with the elicited imitation paradigm (drinking tea and celebrating birthday), were administered.

2.3.1 Encoding

2.3.1.1 Baseline. The infant was encouraged to interact with the event-related props for a fixed time period (90 s). For instance, with the props from the Sleep Event the experimenter said: *"You can play that baby is going to bed using all these things, please show me how you play that the baby is going to bed using all these things"*. The experimenter praised any potentially engaging behavior including behavior not related to the to-be-remembered event. The baseline trial ended when the time had passed, or if the infant became engaged in non-task behavior (e.g., pushing the tray away).

***2.3.1.2 Demonstration*.** Two times in succession the experimenter demonstrated how to produce the events in the specified order with accompanying narration. Because the four events differed, the verbal narration also differed in content (but not in format). The narration for the Sleep Event is given as an example: "*You can play that the baby is going to bed using all these things. Look how I play that the baby is going to bed using all these things! I get the toothbrush from the cup* [step 1], *I put toothpaste on the toothbrush* [step 2], *I brush the baby's teeth* [step 3], *Then I put the baby in the bed* (expected version) / *Then I put the baby under the bed* (unexpected version) [step 4], *and then I put on the duvet* [step 5]. *That is how I play that the baby is going to bed using all these things" ["Let me show you how to do it one more time"]*.

2.3.1.3 Immediate imitation. Immediately after the demonstration, the infant was encouraged to imitate the event for a fixed period of time (90 s) as during baseline. Again, the Sleep Event is used as an example of the instruction given: *"Can you show me how I*

Notice: This is the author's version of a work that was accepted for publication in *Cognitive Development*. A definitive version was subsequently published in *Cognitive Development, 39*, 168-180. DOI: 10.1016/j.cogdev.2016.06.001

played that the baby is going to sleep using all these things?" Note that the experimenter explicitly emphasized that the infant should re-enact the sequence the same way that *the experimenter* had demonstrated the sequence (and hence not necessarily in accordance with the infant's possible script knowledge). Just as in the baseline phase, any potential engaging behavior produced by the infant was praised.

2.3.2 Delayed recall.

Two weeks after the first visit, the infants returned to the lab for the nonverbal delayed recall test. The procedure was identical to the immediate imitation test regarding time period and praising, and the order of the four events was maintained, but the verbal instruction was slightly different. For instance, using the Sleep Event again as an example the experimenter said: *"Last time you were here I showed you, how I play, that the baby is going to bed. Can you show me how I played that the baby is going to bed using all these things?"* Note again, that the infants were explicitly instructed to re-enact the sequences as *the experimenter* had demonstrated to them at the last visit and not just how they would typically tuck the baby into bed.

2.4 Scoring and data reduction

Both sessions were video recorded. From the recordings we derived nonverbal measures of the infants' behavior during baseline, immediate imitation, and delayed recall. First, we scored the number of target actions. Second, and of central interest for the present study, we scored whether or not the infants re-enacted the experimentally manipulated 4th step *as modelled* in each of the four events at both immediate imitation and delayed recall. Remember that the 4th step was either expected or unexpected relative to the scripts in which they were employed.

Two scorers received substantial training based on a scoring manual. The primary scorer (an assistant) then scored all the infants' production of target actions, whereas the secondary scorer (first author) re-scored 25% of the recordings. The interrater agreement was high 96.1 % (Range per child: min = 83.3%; max = 100%).

2.4.1 Number of recalled target actions. The infant received one point for each target action produced correctly. Following Bauer (1992) the infant was only credited for the first occurrence of each target actions that he or she produced. Since all four events were five steps in length, the infant could receive a score in the range of 0-5 for each event.

Notice: This is the author's version of a work that was accepted for publication in *Cognitive Development*. A definitive version was subsequently published in *Cognitive Development, 39*, 168-180. DOI: 10.1016/j.cogdev.2016.06.001

3.0 Results

3.1 Preliminary analyses

We began by analyzing the infants' (re)-enactment of all five target actions involved in events.[1] These analyses were important for two reasons. First, we needed to know whether the basic elicited imitation paradigm worked as planned (i.e., did the infants learn from the instructions, and did they remember across the two-week retention interval?). Second, we wanted to examine whether the two highly familiar events (the Bath and the Sleep event) were indeed more familiar to the infants than the less familiar events (the Car and the Zoo event). In both cases, these analyses served as necessary preconditions for the subsequent and crucial analyses regarding whether the infants re-enacted the expected and unexpected versions of the 4th step as they were demonstrated to them. Descriptive statistics for the mean number of target actions produced for the highly familiar and the less familiar events as well as for each of the four events are presented in Table 1.

3.1.1 The infants' (re)-enactments of target actions. By means of a repeated-measures ANOVA with Familiarity (highly familiar vs. less familiar events) and Time of Test (baseline vs. immediate imitation vs. delayed recall) as within-subjects factors and number of target actions produced as the dependent variable we analyzed whether the infants' (re)-enactments differed with respect to familiarity and the time of test. The analysis revealed a main effect of Familiarity, $F(1, 111) = 504.81$, $p < .001$, $r = 0.91$. The analysis also revealed a main effect of Time of Test, $F(2, 222) = 506.65$, $p < .001$, $r = 0.83$ (see Table 1 for *M*s and *SD*s). Pairwise post hoc analyses using the Bonferroni correction showed that overall the infants (re)-enacted more target actions at the immediate imitation relative to baseline ($p < .001$), that they re-enacted a significant number of target actions across the two-week retention interval, that is, at the delayed recall relative to baseline ($p < .001$), but also that the infants displayed some forgetting across the two-week retention interval, that is, from immediate imitation to delayed recall (see Fig. 3).

Finally, the analyses revealed a significant interaction between Familiarity and Time of Test ($F[2, 222] = 13.04$, $p < .001$, $r = 0.24$), indicating that the effect of familiarity on the

[1] The parenthesis in "(re)-enactment" is used when the analyses concern both baseline and test measures, because whereas the actions sequences at test have been demonstrated to the infants, they have not yet at baseline. Thus, while the prefix "re" makes sense for the test measures, it is somewhat misleading with regards to the baseline measures.

Notice: This is the author's version of a work that was accepted for publication in *Cognitive Development.* A definitive version was subsequently published in *Cognitive Development, 39,* 168-180. DOI: 10.1016/j.cogdev.2016.06.001

number of actions (re)-enacted had more impact on the baseline measures than on the test measures (immediate imitation and delayed recall). In order to qualify this interaction further, additional analyses were conducted. First, by means of three repeated-measures ANOVAs with Familiarity (highly familiar vs. less familiar events) as within-subjects factor and the number of target actions produced at each of the three time points (baseline, immediate imitation, delayed recall) we analyzed whether Familiarity had an effect at each phase of the study. The results revealed that this was the case at all phases of the study as the infants systematically (re)-enacted more target actions with the highly familiar events than with the less familiar events (baseline: $F[1, 111] = 337.89$, $p < .001$, $r = 0.87$; immediate imitation: $F[1, 111] = 163.90$, $p < .001$, $r = 0.77$; delayed recall: $F[1, 111] = 168.46$, $p < .001$, $r = 0.78$, see Table 1 for *M*s and *SD*s). Note that although Familiarity systematically affected the number of actions produced at different time points, the interaction between Familiarity and Time of Test reveals that the Familiarity had a stronger effect on the number of actions enacted by the infants at baseline relative to other time points (see Fig. 3). This is also reflected in the differences in the effect sizes obtained at the three time points: Whereas the effect sizes were almost identical at immediate imitation and delayed recall ($r_{immediate\ imitation} = 0.77$, $r_{delayed\ recall} = 0.78$), the *r*-value was considerably higher at baseline ($r_{baseline} = 0.87$). On average the infants produced more than 2.5 times as many actions at baseline when handed the props used for the highly familiar events relative to the less familiar events. This was in accordance with our predictions and we take this clear and systematic difference to indicate that the infants were indeed more familiar with the two highly familiar events than with the two less familiar events.

Second, we conducted two repeated-measures ANOVAs with Time of Test (baseline vs. immediate imitation vs. delayed recall) as within-subjects factor and the number of target actions produced as the dependent variable (one with the highly familiar events and one with the less familiar events, respectively). For the highly familiar events this analysis revealed a main effect of Time of Test, $F(2, 222) = 212.04$, $p < .001$, $r = 0.70$ (see Table 1 for *M*s and *SD*s). Posthoc analyses using the Bonferroni correction showed that all pairwise comparisons were significant ($p < .001$). Thus, (1) the infants produced more target actions at the immediate imitation relative to baseline; (2) the infants re-enacted more target actions at the delayed recall relative to baseline, and (3) the infants re-enacted fewer target actions at the delayed recall relative to at the immediate imitation. This pattern in results was repeated when looking at the less familiar events, $F(2, 222) = 356.23$, $p < .001$, $r = 0.78$. Likewise,

Notice: This is the author's version of a work that was accepted for publication in *Cognitive Development.* A definitive version was subsequently published in *Cognitive Development*, *39*, 168-180. DOI: 10.1016/j.cogdev.2016.06.001

posthoc analyses using the Bonferroni correction showed that all equivalent pairwise comparisons again were significant ($p < .001$). Thus, the systematic and clear pattern in results across all phases of the study obtained for all events was repeated when considering the highly familiar and the less familiar events alone.

To summarize, the basic elicited imitation paradigm did indeed work out as planned. In addition, before receiving any instructions the infants performed significantly more correct target actions when handed the props related to the highly familiar events (the Bath and the Sleep event) than when offered the props from the less familiar events (the Car and the Zoo event). Note that these results serve as necessary and important preconditions for the crucial analyses, to which we now turn, regarding whether the infants would re-enact the critical 4^{th} step as it had been demonstrated to them across the highly familiar and the less familiar events.

3.2 Was the 4^{th} step re-enacted as demonstrated across conditions and time?

We now turn to the crucial question regarding whether the infants at test re-enacted the 4^{th} step as it had been demonstrated to them. Recall that three factors were manipulated experimentally: (1) Familiarity (highly familiar vs. less familiar event), (2) Expectedness (expected vs. unexpected versions demonstrated), and (3) Time of Test (immediate imitation vs. delayed recall). Because the dependent variable was categorical (whether the 4^{th} step was re-enacted as demonstrated: yes or no) and since more than two factors were potentially relevant, we decided to employ log-linear analysis (Field, 2009). In log-linear analysis all factors, including the dependent variable (i.e., whether the 4^{th} step was re-enacted as demonstrated), are treated as factors (Howel, 1997). Thus, we analyzed the data by means of a four-way hierarchical log-linear analysis with the factors Familiarity (highly familiar vs. less familiar event), Expectedness (expected vs. unexpected version demonstrated), Time (immediate imitation vs. delayed recall), and Outcome (re-enacted the 4^{th} step as demonstrated? yes or no). By means of backward elimination the initial four-way model was reduced to a model including two two-way associations: (1) Time x Outcome, and (2) Expectedness x Outcome, which together retained all effects. The likelihood ratio of this final model was χ^2 (10) = 9.65, $p = .472$. Separate chi-square tests were conducted for Time and Expectedness on Outcome (see Table 2 for frequencies).

Notice: This is the author's version of a work that was accepted for publication in *Cognitive Development.* A definitive version was subsequently published in *Cognitive Development, 39,* 168-180. DOI: 10.1016/j.cogdev.2016.06.001

The analyses revealed a significant association between Expectedness and Outcome, χ^2 (1) = 53.98, $p < .001$. Odds ratios indicated that the odds of re-enacting the 4th step as demonstrated was 2.75 times higher for the infants to whom the expected version of the event had been demonstrated compared to the infants who witnessed the unexpected version of the event. Thus, the infants re-enacted fewer 4th steps as they had been demonstrated to them when they had witnessed the *unexpected* version of the 4th step relative to the expected version of the 4th step. There was also a significant association between Time and Outcome, χ^2 (1) = 39.29, $p < .001$. Odds ratios indicated that the odds of re-enacting the 4th step as demonstrated was 2.36 times higher at immediate imitation compared to delayed recall. Thus, as time went by the infants re-enacted fewer of the 4th steps as they had been demonstrated to them. Finally, and contrary to our expectations, Familiarity of the event had no reliable effect on the infants' re-enactment of the 4th step of the event (χ^2 [1] = 0.11, p = .736; see Fig. 4 for a graphical illustration of the results collapsed across familiarity).[2]

To summarize, whereas the factors Expectedness and Time of Test significantly affected the infants' re-enactment of the 4th step of the event sequences as they had been demonstrated to them, the factor Familiarity had no effect on outcome.

4.0 Discussion

The results from the present study provide clear and systematic evidence that when 24-month-old infants are asked to re-enact previously demonstrated familiar events, then both Expectedness (expected or unexpected versions) and Time of Test (immediate imitation or delayed recall) – but not degree of Familiarity – affected whether they re-enacted the crucial 4th step of the event sequences as it had been demonstrated to them. We begin by discussing the pre-conditional analyses since these served as a necessary prerequisite for the subsequent but central analysis regarding whether the infants re-enacted the 4th step as it had been demonstrated to them.

[2] One might speculate that those of the infants who at baseline actually enacted the expected version of the 4th step without being instructed to do so, would possess a more elaborated (i.e., scripted) understanding of the events relative to those of the infants who did not enact the expected version of the events at baseline. In order to test this possibility we repeated the log-linear analyses but this time the analyses was based only on the infants who at baseline (i.e., prior to any demonstrations) enacted the expected version of the 4th step. However, the results from the analyses from this subsample were actually very similar to the results obtained in the main sample. Due to space limitations the details from the results from the subsample is not given here but can be obtained from the corresponding author.

Notice: This is the author's version of a work that was accepted for publication in *Cognitive Development*. A definitive version was subsequently published in *Cognitive Development*, *39*, 168-180. DOI: 10.1016/j.cogdev.2016.06.001

As evidenced by the baseline measures obtained, the infants enacted without instructions significantly more target actions in relation to the two highly familiar events compared to the two less familiar events, indicating that the two events we assumed to be highly familiar to the 24-month-olds (The Sleep Event and The Bath Event), were indeed more familiar to the infants relative to the two events we assumed to be less familiar to the infants (The Car Event and The Zoo Event).

The results regarding the infants' (re)-enactments of actions across time followed the typical pattern of results seen in elicited imitations studies. Thus, across events, the infants learned from the instructions, they recalled the events above baseline level across the two-week retention interval, but the infants also showed forgetting (e.g., Kingo & Krøjgaard, 2013; Sonne, Kingo, & Krøjgaard, in press). Note that this pattern of results as well as the measures regarding familiarity served as necessary prerequisites by ensuring that the overall design worked as planned. In addition, the infants generally produced more actions with the highly familiar events than with the less familiar events.

We now turn to the crucial question regarding whether the infants re-enacted the 4th step as it had been demonstrated to them and whether their performance was affected by the three factors (Familiarity, Expectedness, and Time of Test) that had been manipulated experimentally. The results from the log-linear analyses revealed that whereas Expectedness and Time of Test had a significant effect on the infants' re-enactment of the 4th step, Familiarity did not. The obtained results were only partially in accordance with our predictions. Based on the fuzzy-trace theory we predicted that the infants would not re-enact the unexpected versions of the familiar events over the two-week retention interval. The obtained results were indeed in this direction, but even more prominent and global than predicted: Both the factors Expectedness and Time of Test did significantly affect whether the infants re-enacted the 4th step of the events as they had been demonstrated to them, but both factors had a *general* impact on the infants' performance (i.e., it was not restricted to specific events). And importantly, whereas the factor Familiarity strongly affected the total number of steps enacted (cf. the preliminary analyses), Familiarity turned out to have no influence on whether the 4th step was re-enacted or not. The latter results came as a surprise (see below).

In accordance with the fuzzy-trace theory the obtained results show that the infants re-enacted fewer of the unexpected 4th steps of the events as they had been demonstrated to them (i.e., the verbatim). Because the unexpected incidents clearly violated the scripts, and

Notice: This is the author's version of a work that was accepted for publication in *Cognitive Development.* A definitive version was subsequently published in *Cognitive Development, 39,* 168-180. DOI: 10.1016/j.cogdev.2016.06.001

since this aspect was the only difference between the unexpected and the expected versions of the event sequences demonstrated, we attribute this difference in results to scripted knowledge possessed by the infants. The results also revealed that the infants' re-enactments changed reliably across time: Whereas the infants at the immediate imitation overall tended to re-enact the 4th step as it had been demonstrated to them, this pattern of results became significantly less pronounced at the delayed test, probably due to forgetting of verbatim information. Thus, the results seem to suggest that scripted knowledge did indeed affect the infants' re-enactment of the event sequences evidenced by the fact that the infants re-enact reliably fewer of the unexpected version of the events.

Meanwhile, the obtained results appear to be in sharp contrast to the predictions from the distinctiveness theory by which we, as a contrasting hypothesis, predicted that the unexpected versions would standout and prevail due to their distinctiveness. This is interesting because the explicit instruction provided by the experimenter at the tests unambiguously requested the infants to re-enact the event as *the experimenter* had demonstrated the event to them and *not* as the event may typically occur in daily life. Thus, the instruction explicitly encouraged the infants to re-enact the distinct (and at times unexpected) version previously demonstrated to them. Although the design in this respect may have been said to give preference to the distinctiveness theory, the results provide compelling evidence against the distinctiveness theory under the given circumstances. Note further, that the negative results regarding distinctiveness theory cannot be explained by reference to generally poor memory of the events by the infants. The preliminary analyses showed convincingly (see Table 1 and Figure 3) that overall the infants' (re)-enactments followed the typical pattern of results obtained in elicited imitation studies: the infants learned from the instructions; they remembered across the two-week retention interval, but also displayed some forgetting. Besides, and more important, the unexpected 4th steps of the events, which ought to stand out according to distinctiveness theory, were *generally* re-enacted as demonstrated to a significantly less degree than the expected versions of the events, that is, regardless of the time of test.

Why did the factor Familiarity have no effect on the infants' re-enactments of the 4th step? Note that preliminary analyses revealed that the highly familiar events were indeed more familiar to the infants than the less familiar events. However, although the infants were more familiar with the highly familiar events than with the less familiar events, the infants were presumably somewhat familiar with *both* types of events as indicated by the relatively

Notice: This is the author's version of a work that was accepted for publication in *Cognitive Development.* A definitive version was subsequently published in *Cognitive Development, 39,* 168-180. DOI: 10.1016/j.cogdev.2016.06.001

high average number of actions produced at baseline, before any instructions were given. This interpretation is supported by the baseline measures obtained with familiar events in the previously cited study by Bauer and Travis (1993, Exp. 2). Both studies tested 24-month-old infants on 5-step familiar events, and the magnitude of number of actions produced at baseline were equivalent. We speculate that the average relatively high level of pre-test knowledge that the infants seem to have brought to the lab in the present study regarding each of the four sets of props before any instruction were given, may have been sufficient to allow for building up expectations regarding what to experience (i.e., the expected versions). Hence, when the documented reliable difference in prior knowledge with regards to the highly familiar and the less familiar events had no impact on the infants' re-enactment of the 4th step as it had been demonstrated to them, it may have been because possessing *some* scripted knowledge (which the infants on average had for all events) may be relatively more important in the present context than the significant difference in familiarity between the highly familiar and the less familiar events.

The results from the present study showed that scripted knowledge in the form of expectedness played a role in 24-month-old infants' re-enactment of the events. The obtained results are in accordance with evidence from studies investigating young children by verbal measures showing that scripted knowledge tends to guide their recall of familiar events (e.g., Nelson & Gruendel, 1981, 1986; Slackman & Nelson, 1984). The role of distinctiveness in memory has hitherto been unexplored in infancy (Howe, 2006). The results from the present study seem to suggest that distinctive incidents are not well preserved in infant memory, that is, at least not under the circumstances tested here.

When considering episodic memory, the prevailing view is that episodic memory is a late developing achievement in the ontogenesis (Tulving, 1984). The results obtained in the present study may suggest that research on how scripted knowledge can interfere with the maintenance of the distinct memory trace (the verbatim trace) may shed new light on some of the mechanisms responsible for the apparent late development of the episodic memories. In this regard the results from the present study may suggest that if the infant has developed a familiar script, then atypical or unexpected incidents of events categorized as belonging to the script become more difficult to recall. However, alternative interpretations are possible too. For instance, it may be the case that the infants did in fact remember the unexpected version, but chose to re-enact the expected version – maybe in order to achieve a desired end-state (i.e. wanting to see the dull being put *in* the bed rather than under it). Evidence

Notice: This is the author's version of a work that was accepted for publication in *Cognitive Development.* A definitive version was subsequently published in *Cognitive Development, 39*, 168-180. DOI: 10.1016/j.cogdev.2016.06.001

from other studies show that the achievement of a preferred end-state may at times be more important for infants than re-enacting a specific action sequence verbatim (e.g., Brugger, Lariviere, Mumme, & Bushnell, 2007; Gergely, Bekkering, & Király, 2002). One way to examine whether the infants actually possess a representation of the unexpected version while producing the *expected* version would be to measure event-related potentials during the tests.[3] While we acknowledge that we cannot confidently rule out this alternative interpretation given the available data, we tend to prefer the interpretation offered here: that when the infant has developed a familiar script, then atypical or unexpected incidents of events categorized as belonging to the script may become more difficult to recall. The reasoning is that this pattern was consistent across events, and possibly more important, that the infants in the present study were explicitly instructed to re-enact the modelled sequence and not just the end-state.

The present findings are also relevant when attempting to understand how script deviations are represented in memory, and how unusual information is used to update or modify young children's script knowledge (cf. Hudson et al., 1992). Thus, in this light the present study may be seen as an attempt to re-vitalize the almost abandoned trend of studying the role of scripts on memory in infancy.

While the present study contributes to the understanding of how scripted knowledge affects memory of events, it at the same time raises new questions. First, do infants possess representations of the unexpected version, while producing the expected version at test? Second, (providing our interpretation is correct) 'how much' familiarity with a given event type is necessary and sufficient to establish a gist that over time may overwrite an unexpected incident of the given event type? Third, will scripted knowledge of familiar events tend to overwrite the memory trace of unexpected events in even younger infants than the 24-month-olds examined in the present study? Fourth, how would the length of the retention interval affect the possible impact of scripted knowledge across time? Would the results regarding Time of Test have been the same if the retention interval had been reduced to for instance a single week? And for how long will distinct unexpected incidents be available in memory before presumably being absorbed by the familiar script? We find such questions both interesting and important and hope that the present study will encourage researchers to re-vitalize this almost abandoned tradition of examining memories of scripted

[3] We thank the reviewer for suggesting this alternative interpretation and a possible solution for deciding between the two interpretations.

events in infancy by means of the elicited imitation paradigm. Evidence from such studies are likely to stimulate not only this specific area, but may also contribute to our knowledge on the development of the other side of the coin, that is, the development of memory for specific events.

Acknowledgement

This research was supported by the Danish National Research Foundation (Grant DNRF89), VELUX FONDEN (#10386) and MIND*Lab*. First, we are grateful to Robyn Fivush and Patricia Bauer for their constructive comments regarding the four scripts used in the present study. Second, we would like to thank to Bo Sommerlund and Sarah van Mastrigt for statistical assistance and Trine Ravn for scoring the data. Finally, we would like to thank the children who participated in this study and their parents.

Notice: This is the author's version of a work that was accepted for publication in *Cognitive Development.* A definitive version was subsequently published in *Cognitive Development, 39,* 168-180. DOI: 10.1016/j.cogdev.2016.06.001

References

Bauer, P.J. (1992). Holding it all together: How enabling relations facilitate young children's event recall. *Cognitive Development, 7,* 1-28. DOI: 10.1016/0885-2014(92)90002-9

Bauer, P. J., & Mandler, J. M. (1989). One thing follows another: Effect of temporal structure on 1- to 2-year-olds' recall of events. *Developmental Psychology, 25,* 197-206. DOI.org/10.1037/0012-1649.25.2.197

Bauer, P. J., & Mandler, J. M. (1992). Putting the horse before the cart: The use of temporal order in recall of events by one-year-old children. *Developmental Psychology, 28*, 441-452. DOI.org/10.1037/0012-1649.28.3.441

Bauer, P. J., & Thal, D. J. (1990). Scripts or scraps: Reconsidering the development of sequential understanding. *Journal of Experimental Child Psychology, 50,* 287-304. DOI: 10.1016/0022-0965(90)90043-8

Bauer, P. J., & Travis, L. L. (1993). The fabric of an event: Different sources of temporal invariance differentially affect 24-month-olds' recall. *Cognitive Development, 8*, 319-341. DOI: 10.1016/S0885-2014(93)80004-D

Bauer, P. J., & Shore, C. M. (1987). Making a memorable event: Effects of familiarity and organization on young children's recall of action sequences. *Cognitive Development, 2,* 327-338. DOI: 10.1016/S0885-2014(87)80011-4

Brainerd, C. J., & Reyna, V. F. (2001). Fuzzy-trace theory: Dual processes in memory, reasoning, and cognitive neuroscience. *Advances in Child Development and Behavior, 28,* 41-100. PMID: 11605365

Brainerd, C. J., & Reyna, V. F., & Kneer, R. (1995). False-recognition reversal: When similarity is distinctive. *Journal of Memory and Language, 34,* 157-185. DOI: 10.1006/jmla.1995.1008

Brainerd, C. J., Reyna, V. F., Wright, R., & Mojardin, A. H., (2003). Recollection rejection: False-memory editing in children and adults. *Psychological Review*, *110,*762-784. DOI.org/10.1037/0033-295X.110.4.762

Brugger, A., Lariviere, L.A., Mumme, D.L., & Bushnell, E.W. (2007). Doing the right thing: Infants' selection of actions to imitate from observed event sequences. *Child Development, 78,* 806-824. DOI: 10.1111/j.1467-8624.2007.01034.x

Dahl, J. J., Sonne, T., Kingo, O. S., & P. Krøjgaard (2013). On the development of episodic memory: Two basic questions. *Nordic Psychology*, *56,*189-207. DOI: 10.1080/19012276.2013.807661

Davidson, D., & Hoe, S. (1993). Children's recall and recognition memory for typical and atypical actions in script-based stories. *Journal of Experimental Child Psychology, 55,* 104-126. DOI: 10.1006/jecp.1993.1005

Field, A. (2009). *Discovering statistics using SPSS.* (3rd Ed.). London: Sage.

Fivush, R. (1984). Learning about school: The development of kindergartners' school scripts. *Child Development*, *55,* 188-201. DOI: 10.1111/j.1467-8624.1984.tb00412.x

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

Fivush, R., & Mandler, J. M. (1985). Developmental changes in the understanding of temporal sequence. *Child Development, 56,* 1437-1446. DOI: 10.1111/j.1467-8624.1985.tb00209.x

Gergely, G., Bekkering, H., & Király, I. (2002). Developmental psychology: Rational imitation in preverbal infants. *Nature*, *415*, 755. DOI:10.1038/415755a

Howe, M. L. (2006). Distinctiveness effect in children's memory. In R. R. Hunt & J. B. Worthen (Eds.), *Distinctiveness and memory* (pp. 237-257). Oxford University Press, Inc., New York.

Howel, D. C. (1997*). Statistical methods for psychology, 4th ed.* Belmont: Duxbury Press

Hudson, J. A. (1990). Constructive processes in children's event memory. *Developmental Psychology*, *26,* 180-187. DOI.org/10.1037/0012-1649.26.2.180

Hudson, J. A., Fivush, R., & Kuebli, J. (1992). Scripts and episodes: The development of event memory. *Applied Cognitive Psychology*, *6,* 483-505. DOI: 10.1002/acp.2350060604

Hudson, J. A., & Mayhew, E. M. Y. (2009). The development of memory for recurring events. In M. L. Courage & N. Cowan (Eds.), *The development of memory in infancy and childhood* (pp. 69-91). Psychology Press. Taylor & Francis Group. Hove and New York.

Hudson, J. A., & Nelson, K. (1983). Effects of script structure on children's story recall. *Developmental Psychology, 19,* 625-635. DOI.org/10.1037/0012-1649.19.4.625

Hudson, J. A., & Nelson, K. (1986). Repeated encounters of a similar kind: Effects of familiarity on children's autobiographical memory. *Cognitive Development, 1*, 253-271. DOI: 10.1016/S0885-2014(86)80004-1

Notice: This is the author's version of a work that was accepted for publication in *Cognitive Development*. A definitive version was subsequently published in *Cognitive Development, 39*, 168-180. DOI: 10.1016/j.cogdev.2016.06.001

Hunt, R. R. (2006). The concept of distinctiveness in memory research. In R. R. Hunt & J. B. Worthen (Eds.), *Distinctiveness and memory* (pp. 3-25). Oxford University Press, Inc., New York.

Hunt, R. R., & Worthen, J. B. (Eds.) (2006). *Distinctiveness and memory*. Oxford University Press, Inc., New York.

Kingo, O. S., & Krøjgaard, P. (2013). Eighteen-month-old infants generalize to analogue props across a two weeks retention interval in an elicited imitation paradigm. *Child Development Research, 2013,* 1-11. http://dx.doi.org/10.1155/2013/786862

Krøjgaard, P. (2005). Spædbarnsforskningens relevans for almenpsykologien [The releveance of infant research in relation to general psychology. (Target article with open peer commentaries)]. *Bulletin fra Antropologisk Psykologi, 15,* 2-27.

Kuebli, J., & Fivush, R. (1994). Children's representations and recall of event alternatives. *Journal of Experimental Child Psychology, 58,* 25-45. DOI: 10.1006/jecp.1994.1024

Mandler, J. M. (1998). Babies think before they speak. *Human Development, 41,* 116-126. DOI: 10.1159/000022572

Nelson, K. (1978). How young children represent knowledge of their world in and out of language. In R. S. Siegler (Ed.), *Children's thinking: What develops?* (pp. 255-273). Hillsdale, NJ: Erlbaum.

Nelson, K. (1979). Children's long-term memory for routine events. Paper presented at the Annual Meeting of the American Psychological Association (87th, New York, NY, September 1-5. ED178222

Nelson, K. (1988). The ontogeny of memory for real events. In U. Niesser (Ed.), *Remembering reconsidered: Ecological and traditional approaches to the study of memory* (pp. 244-276). New York: Cambridge University Press.

Nelson, K. & Gruendel, J. M. (1981). Generalized event representations: Basic building blocks of cognitive development. In M. E. Lamb & A. L. Brown (Eds.), *Advances in Developmental Psychology, Vol. 1,* (pp. 131-158)

Nelson, K., & Gruendel, J. M. (1986). Children's scripts. In K. Nelson (Ed.), *Event knowledge: Structure and function in development* (pp. 21-46). Hillsdale, NJ: Erlbaum.

O'Connell, B. O., & Gerard, A. B. (1985). Scripts and scraps: The development of sequential understanding. *Child Development, 56,* 671-681. DOI: 10.1111/j.1467-8624.1985.tb00141.x

Piaget, J. & Inhelder, B. (1973). *Memory and intelligence.* New York: Basic Books.

Shank, R., & Abelson, R. (1977). *Scripts, plans, goals and understanding*. Hillsdale, NJ: Erlbaum.

Slackman, E., & Nelson, K. (1984). Acquistition of an unfamiliar script in story form by young children. *Child Development, 55,* 329-340. DOI: 10.2307/1129946

Sonne, T., Kingo, O.S., & Krøjgaard, P. (in press). Empty looks or paying attention? Exploring infants' visual behavior during encoding of an elicited imitation task. *Infancy*. DOI: 10.1111/infa.12141

Tulving, E. (1984). Précis of elements of episodic memory. *The Behavioral and Brain Science*, *7,* 223-268. DOI: 10.1017/S0140525X0004440X

Notice: This is the author's version of a work that was accepted for publication in *Cognitive Development*. A definitive version was subsequently published in *Cognitive Development, 39*, 168-180. DOI: 10.1016/j.cogdev.2016.06.001

Table 1

Means and Standard Deviations for the number of different Target Actions for the Highly Familiar (Bed and Bath events combined) and the Less Familiar (Car and Zoo events combined) events. The maximum score for target actions were 5 points for each event.

	Actions (Re)enacted					
	Baseline		Immediate Imitation		Delayed Recall	
Events	*M*	*(SD)*	*M*	*(SD)*	*M*	*(SD)*
Highly Familiar	2.57	(0.83)	4.28	(0.61)	3.74	(0.81)
The Bed Event	3.16	(1.14)	4.53	(0.83)	4.02	(1.11)
The Bath Event	1.97	(1.04)	4.04	(0.83)	3.46	(1.06)
Less Familiar	0.95	(0.55)	3.24	(0.83)	2.51	(0.79)
The Car Event	1.01	(0.80)	3.41	(1.08)	2.85	(1.10)
The Zoo Event	0.88	(0.69)	3.06	(1.20)	2.18	(1.11)

Notice: This is the author's version of a work that was accepted for publication in *Cognitive Development.* A definitive version was subsequently published in *Cognitive Development*, *39*, 168-180. DOI: 10.1016/j.cogdev.2016.06.001

Table 2

Complete sample including all infants. Frequencies and overall percentages for the observed outcomes of the infants' re-enactment of the 4th step at Test across Familiarity, Expectedness, and Time of Test are displayed.

				Observed	
Familiarity	**Expectedness**	**Time of Test**	**4th step re-enacted as shown?**	Count[a]	%
Highly Familiar	Expected	Immediate	Yes	84	9.4
			No	28	3.1
		Delayed	Yes	65	7.3
			No	47	5.2
	UnExpected	Immediate	Yes	72	8.0
			No	40	4.5
		Delayed	Yes	35	3.9
			No	77	8.6
Less Familiar	Expected	Immediate	Yes	89	9.9
			No	23	2.6
		Delayed	Yes	70	7.8
			No	42	4.7
	UnExpected	Immediate	Yes	55	6.1
			No	57	6.4
		Delayed	Yes	37	4.1
			No	75	8.4
Total				896	100

a. Because all infants were shown two highly familiar (one expected and one unexpected) and two less familiar (one expected and one unexpected) at two time points (immediate and delayed test), the total number of observations were 112 (participants) x 8 (conditions) = 896 observations.

Notice: This is the author's version of a work that was accepted for publication in *Cognitive Development*. A definitive version was subsequently published in *Cognitive Development, 39*, 168-180. DOI: 10.1016/j.cogdev.2016.06.001

Figure Captions

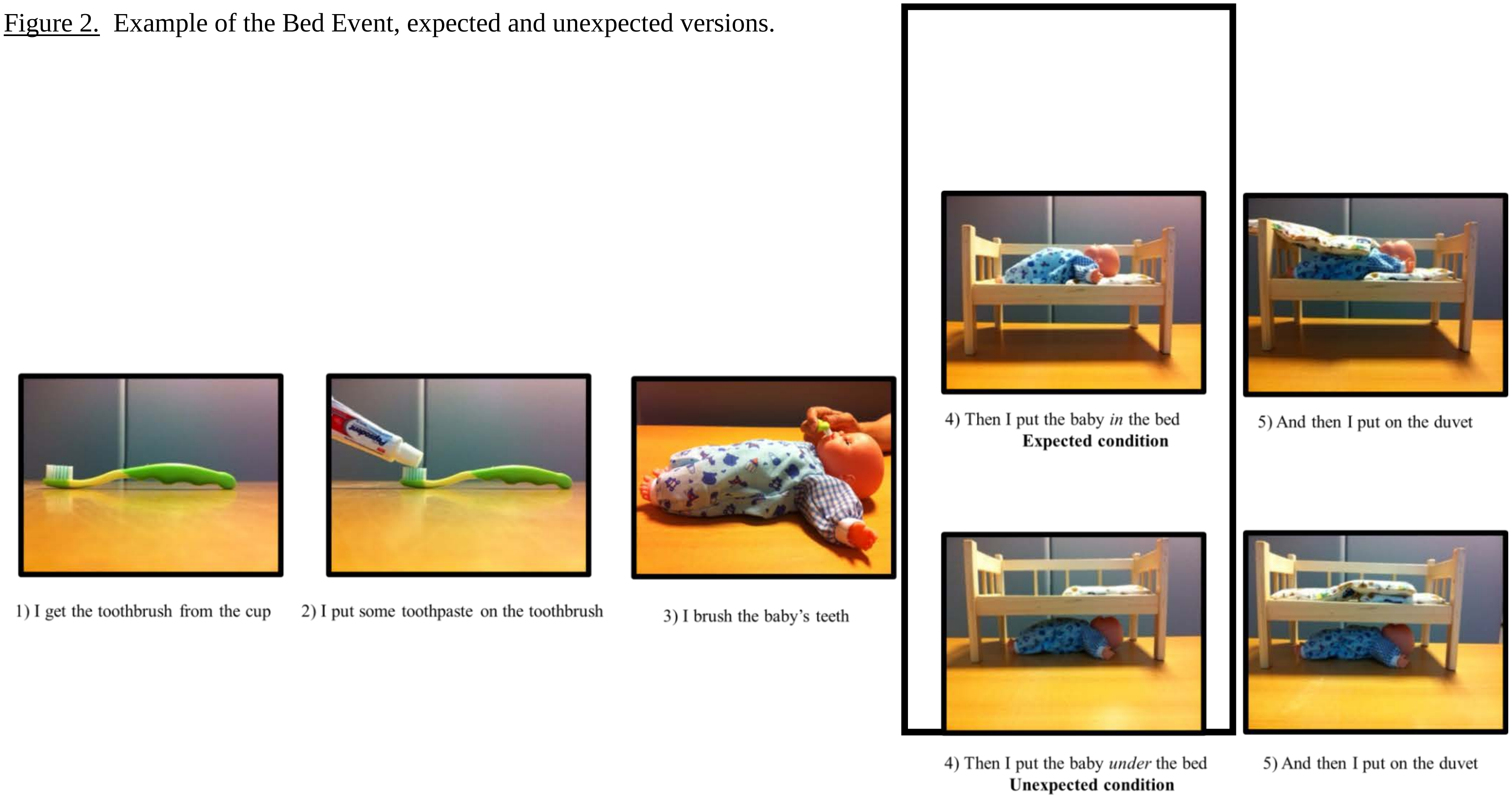
Figure 1. Pictures of the four events and the props used to re-enact the four events as they were demonstrated to the infants on the tray (seen from the child's perspective) at Baseline, Demonstration, Immediate Imitation, and Delayed Recall.

Figure 1.

Notice: This is the author's version of a work that was accepted for publication in *Cognitive Development.* A definitive version was subsequently published in *Cognitive Development*, *39*, 168-180. DOI: 10.1016/j.cogdev.2016.06.001

<u>Figur 2.</u>

<u>Figure 2.</u> Example of the Bed Event, expected and unexpected versions.



Notice: This is the author's version of a work that was accepted for publication in *Cognitive Development.* A definitive version was subsequently published in Cognitive Development, 39, 168-180. DOI: 10.1016/j.cogdev.2016.06.001

Figure 3

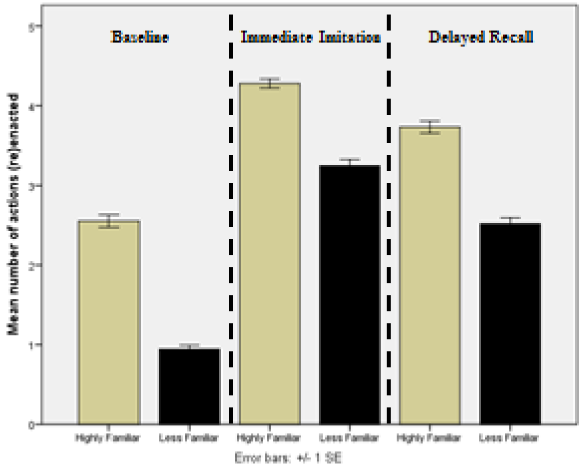


Figure 3. Graphical presentation of the mean number of target actions at Baseline, Immediate Imitation, and Delayed Recall, (re)enacted across Highly Familiar and Less familiar events.

Figure 4.

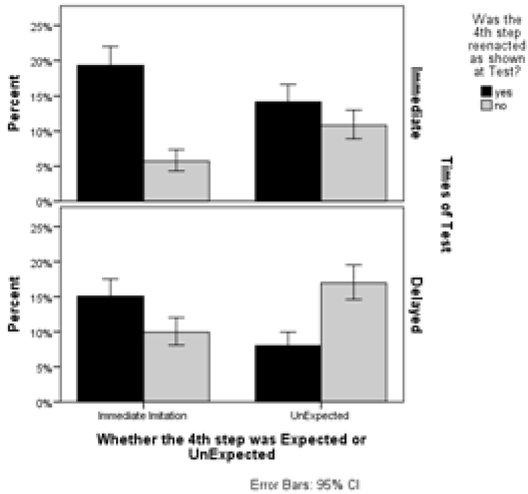


Figure 4. Graphical representation of whether the infants reenacting the events as they had been demonstrated to them across the two of the manipulated factors that revealed a main effect: Expectedness, and Time of Test.