Research Article

Eighteen-Month-Old Infants Generalize to Analog Props across a Two-Week Retention Interval in an Elicited Imitation Paradigm

Osman S. Kingo and Peter Krøjgaard

Center on Autobiographical Memory Research, Department of Psychology and Behavioural Sciences and MINDLab, Aarhus University, Bartholins Allé 9, 8000 Aarhus C, Denmark

Correspondence should be addressed to Osman S. Kingo; osman@psy.au.dk

Received 29 November 2012; Revised 30 January 2013; Accepted 1 February 2013

Academic Editor: Jeffrey W. Fagen

Copyright © 2013 O. S. Kingo and P. Krøjgaard. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We report a generalization experiment in which 72 18-month-old infants were tested in the elicited imitation paradigm. Two questions were addressed: (1) whether infants were able to generalize to differently looking (shape and color changes) but functionally equivalent props and (2) whether narrative support at both encoding and retrieval would facilitate memory. The results revealed that the 18-month-old infants were indeed capable of generalizing to differently looking but functionally equivalent props across a retention interval of two weeks. However, contrary to expectations, narrative support did not facilitate memory or generalization.

1. Introduction

Compared to other species, human beings seem to excel in the ability to transfer knowledge from one domain to another [1]. This ability is a basic requirement behind the idea of having formalized educational systems like schools where children are supposed to acquire knowledge in many different versatile domains like reading, writing, and math in order to be able to apply such knowledge to different domains outside the school. Knowledge transfer is also a crucial component in creativity and analogical problem solving (e.g., [2]). This raises the question regarding when and how infants and children begin to be able to generalize knowledge from one domain to another. However, in order to transfer knowledge from one domain to another one will have to remember what has been learned.

Infant memory has been investigated by means of at least three different paradigms: (1) visual habituation, (2) conjugate reinforcement, and (3) deferred and elicited imitation [3–6]. Studies from the latter two paradigms especially have addressed to what extent infants are capable of transferring knowledge across domains. Results from these two paradigms are treated separately in the following.

In an extensive series of studies using the conjugate reinforcement paradigm, Rovee-Collier and her collaborators

have, for instance, systematically investigated how changes between the encoding setting and the retrieval setting influence infants' memory abilities across age (for a recent review, see [6]). The basic conjugate reinforcement design involves 2–6-month-old infants placed under a mobile with a ribbon attached to his or her leg. After having established a baseline measure of the number of kicks per minute with the ribbon loosely attached to the fence of the bed, the ribbon is attached to the mobile in such a manner that whenever the infant kicks the mobile moves for a nine minutes learning sequence. This is followed by an encoding test: the ribbon is moved from the mobile and back to the fence of the bed. If the infant for a three-minute period produces significantly more kicks compared to baseline, it is taken as evidence that the infant has learned an association between kicking and having the mobile move. After a retention period, infants have been brought back for further testing in order to see under which circumstances memory is preserved. Results from such studies have shown that three-month-old infants remember what they have learned across a retention interval of one week [7] and that six-month-old infants preserve memory over retention intervals of two weeks [8].

However, young infants' memory abilities in the conjugate reinforcement paradigm have been shown to be highly sensitive to changes between the learning conditions and

the retrieval conditions. If for instance more than one of the mobile's objects had been changed between learning and retrieval, then memory was disrupted in 2- and 3-month-old infants [9]. Changes in the context between learning and retrieval have also been shown to reduce memory performance in 6-month-old infants (e.g., [10]). However, reactivation studies where a forgotten memory trace is reinstated by means of cuing have shown that infants from around the age of nine months become better at generalizing their memories across different contexts (e.g., [11]).

Infants' abilities regarding transferring knowledge across domains have also been investigated by means of the deferred or elicited imitation paradigms. In deferred or elicited imitation paradigms, an adult models a distinct action, or action sequence, and after a delay (from minutes to months), the infant is given the opportunity to reproduce the to-be-remembered action or sequence of actions [3, 12]. The dependent variable is the number of correctly reproduced *steps* involved in the to-be-remembered action sequence, and if more than one step is involved, then also the number of correctly ordered *pairs* of steps reproduced. Deferred and elicited imitation differ with regard to the employment of encoding tests presented immediately after the demonstration. Designs using encoding tests are usually referred to as elicited imitation, whereas studies leaving out encoding tests are typically called deferred imitation [3]. Moreover, in elicited imitation studies, the experimenter typically accompanies the demonstration by narrative support (e.g., [13]). The deferred and elicited imitation paradigms have both proven to be powerful tools for investigating infant memory [12].

Results from the deferred imitation paradigm have revealed that 14-month-old infants were able to reproduce the to-be-remembered actions across a 10-minute retention interval even though the demonstration and the test took place in two different contexts [14]. In a subsequent study using the same paradigm, 12-month-old infants showed recall memory across different contexts for delays up to four weeks [15].

When using imitation paradigms, the learning context can differ from the test context in other respects than with regard to the specific location. Recent evidence suggests that infants are capable of reproducing action sequences that have been demonstrated via 2D DVD presentations or 2D books, instead of the typical live, face-to-face demonstrations. For example, Simcock et al. [16] showed that 18- and 24-month-old infants successfully reproduced three-step action sequences that had been presented via 2D DVD recordings and 2D books across a retention interval of 10 minutes. Infants in both age groups reliably reproduced the action sequences regardless of whether they had received "full narration" (where the experimenter supported the demonstration by accompanying the demonstration with words that directly referred to the action sequence) or "empty narration" (where the experimenter provided narration but without referring explicitly to the action sequences). However, a systematic manipulation of the specificity of prompts used at the test revealed that infants who received explicit prompts ("Show me how to make a rattle") fared significantly better than infants who received nondirective prompts ("Show me how

to make something") [16]. In a very recent study by Brito et al. [17], the authors extended the results obtained by Simcock and colleagues [16] by showing that action sequences learned via 2D presentations was preserved across retention interval of two weeks for 18-month-old infants and across four weeks for 24-month-old infants [17].

Are infants capable of recalling the to-be-remembered actions in an imitation paradigm if the props used for the action sequences are changed between demonstration and test to differently looking (e.g., changing size, shape, and color), but functionally equivalent props? Using the deferred imitation paradigm, 14-month-old infants were able to recall the demonstrated events across a 10-minute retention interval even though the props used for the test were different with respect to size and color [14]. Herbert and Hayne [18, Exp. 1] investigated 18- and 24-month-old infants' ability to generalize to differently looking, but functionally equivalent props over a 24 h delay. Only the 24-month-old infants were able to recall the to-be-remembered actions across the 24 h delay. By means of the elicited imitation paradigm, Bauer and Dow [19] showed that both 16- and 20-month-old infants were able to recall the to-be-remembered actions across a retention interval of one week even though the props had been changed from demonstration to test to differently looking but functionally equivalent ones. Thus, at the time of writing, a retention interval of one week is the longest obtained so far by which infants have been able to generalize to differently looking but functionally equivalent props.

In the present study, we attempted to investigate whether 18-month-old infants would be able to generalize to differently looking but functionally equivalent props across a retention interval of two full weeks while systematically manipulating a factor that might be of crucial importance for encoding and retrieval, that is, the amount of narrative support from the experimenter.

Why should narrative support influence encoding and/or retrieval? A large number of studies have shown that maternal reminiscing style have substantial influence on children's ability to report episodes from their own lives. Mothers entertaining an elaborative maternal reminiscence style tend to have children who report earlier and more elaborated autobiographical memories (e.g., [20–22]). Some studies have investigated the possible effect of talking about the to-be-remembered material, while the to-be-remembered material is being encoded (e.g., [23, 24]). For instance, Tessler and Nelson [24, study 1] had 3–3.5-year-old children visiting a museum together with their mothers. Half of the mothers were instructed to interact normally with the children during the visit, whereas the other half were asked only to reply to specific questions from the children, but not to initiate any questions or elaborate on the children's responses. One week later, the children were interviewed about what they remembered from the visit. None of the children showed any recall of items from the museum they had seen, but not talked about [24]. Thus, having an adult talking about and elaborating on the event as it unfolds seems to facilitate the child's ability to remember the event at a later point in time.

However, these studies were conducted with preschool children. What if the children are younger and less mature

language communicators? Will narrative support still facilitate retrieval? Only few studies have systematically investigated the potential effects of narrative support during demonstration and test (e.g., [25]), and even fewer studies have investigated the possible interaction of narrative support on generalization tasks.

Only two studies of the latter kind have been conducted to date: in the previously mentioned generalization study by Herbert and Hayne [18, Exp. 2], the experimenter applied unique (nonsense) verbal labels at both demonstration and test in order to see whether this would facilitate generalization. The results revealed that applying verbal labels only helped the 24-month-olds but not the 18-month-old infants to generalize across differently looking props across the 24 h retention interval. And in a recent study by Herbert [26] employing a 10-minute retention interval, 12- and 15-month-old infants were given either Empty narration (e.g., “Look at this”) or Full narration (e.g., “Look. A puppet”) at both demonstration and test (i.e., not counterbalanced). The infants in both age groups were most successful when given Full narration compared to Empty narration [26]. Thus, the recent study by Herbert [26] was the first to demonstrate a positive influence of language cues on a generalization task for infants below 24 months of age.

Hayne and Herbert [27] did not investigate generalization but conducted the first study in which narrative support at both demonstration and test across a four-week retention interval was investigated systematically with 18-month-old infants. The infants received either “Empty” or “Full” narration at both demonstration and test. Over two experiments, all possible four combinations for the demonstration and test were present: FF (Full, Full), EF (Empty, Full), FE (Full, Empty), and EE (Empty, Empty). The results revealed that infants receiving Full narration were superior with regard to memory compared to the infants receiving Empty narration. The effect of Full narrative support was only present at test, but not at encoding. Furthermore, infants in the EF condition fared better compared to the infants in the FE condition. It should be mentioned that in the previously mentioned study by Bauer and Dow [19] in which generalization to differently looking but functionally equivalent props was demonstrated across the hitherto longest retention interval, the experimenter invariably employed narration at *both* demonstration and test.

As evidenced above, several studies have documented that narrative support does facilitate memory performance in imitation tasks in infancy. However, recent evidence with 15-month-old infants suggests that this is not always the case. In an imitation study by Zack et al. [28], narrative support did not help the 15-month-old infants to transfer their experiences from 2D demonstrations to 3D tests (or vice versa).

To summarize, older infants appear to be able to generalize to functionally equivalent but differently looking props provided that the retention interval is relatively short (e.g., one week or shorter). Moreover, although narrative support seems to facilitate retrieval in children, the effect on infants appears to be mixed. Based on these findings we decided to investigate 18-month-old infants’ ability to generalize to

functionally equivalent but differently looking props across a retention interval of two weeks, while systematically manipulating the narrative support during demonstration as well as at the delayed test. Given the limited available evidence we tentatively hypothesized that the 18-month-old infants might be able to generalize across prop changes, but only when provided with narrative support at retrieval.

2. Method

2.1. Participants. A total of 72 18-month-old infants (38 girls and 34 boys; $M_{\text{age, total}} = 17$ months 29 days; range 17 months and 20 days to 18 months and 10 days) were randomly assigned to one of four conditions (Full-Full: $n = 18$ (9 girls, 9 boys); $M_{\text{age}} = 17$ months and 29 days; range: 17 months and 21 days to 18 months and 9 days; Full-Empty: $n = 18$ (11 girls, 7 boys); $M_{\text{age}} = 17$ months and 29 days; range: 17 months and 18 days to 18 months and 9 days; Empty-Full: $n = 18$ (8 girls, 10 boys); $M_{\text{age}} = 18$ months and 2 days; range: 17 months and 21 days to 18 months and 12 days; Empty-Empty: $n = 18$ (10 girls, 8 boys); $M_{\text{age}} = 17$ months and 28 days; range: 17 months and 18 days to 18 months and 9 days). Only infants who completed tests for all five props were included (Four of the five props used were replicas of props developed and used by Patricia Bauer. We would like to thank Patricia Bauer for kindly giving us access to detailed information on these props. A sixth prop, the Slide, was also used. However, since all infants had severe difficulties (re)producing any of the steps related to the Slide—probably because (a) the Slide turned out to be more difficult to handle motor-wise and (b) because it probably did not afford the action steps involved—it was excluded from the data analysis. Please note that the Slide that turned out not to work was an invention of our own and not an invention from Bauer’s lab.). An additional 25 infants were tested but excluded from the analysis (10 due to fussiness, 1 due to experimenter error, 1 due to parent interference, 1 due to shyness, 1 due to technical problems, and 11 because of missing data in at least one subtest). The infants were recruited through access to a birth register. All participating infants were full term and healthy. Besides receiving a small participant gift, no compensation was offered.

2.2. Apparatus. The infants were seated in a high chair at a 180 cm wide and 90 cm deep table in front of the experimenter. Two HDD cameras recorded the infants and the experimenter for later scoring. The props were presented on a 55 × 37 cm rectangularly shaped wooden tray equipped with a 3 cm high edge. Five different 3-step props were used: A Gong, a Jumping Jack, a Spinner, a Shaker, and a Windmill (see Figure 1).

Each prop had been made in two versions: a target version and an analog version. The analog versions were functionally equivalent to the target versions but differed from the target versions by having different colors (e.g., white instead of blue) and different shapes (e.g., square base instead of round base), whereas sizes were kept approximately constant. All props were constructed specifically for the study and were not commercially available.

FIGURE 1: Pictures of props used in the study.

The Gong consisted of a 24.5×24.5 cm square shaped, red base with two 16 cm tall, red posts with a red bar attached to one of the posts; a green square shaped 9×9 cm metal plate, and a 21 cm long wooden hammer. The analog version of the Gong had a 6-edged base (maximum distance between edges = 24 cm), and the base, the posts, and the bar were white. The metal plate was 6-edged and painted red (maximum distance between edges = 10 cm), and the club was yellow and red, made of wood, and 16 cm long. The three-step sequence to be remembered for the Gong was the following: first, the bar is flipped over in order to produce a cross bar (step 1). Then the plate is mounted on the cross bar (step 2), and finally you ring the Gong by hitting the plate with the hammer (step 3).

The Jumping Jack consisted of a blue, circular shaped base (diameter = 14 cm) with a center-placed hole; a yellow 27 cm long wooden stick with a strip of Velcro attached on the top part; and a Jumping Jack yellow and brown tiger (maximum $h \times w = 16 \times 16$ cm) with Velcro mounted on the back side. The analog version of the Jumping Jack had a red rectangular shaped base ($l \times w = 10 \times 14$ cm), a grey 27 cm long stick with Velcro, and a green Jumping Jack crocodile (maximum 15 cm tall and 12 cm wide) with Velcro mounted on the back. The three-step sequence to be remembered for the Jumping Jack was as follows: the stick is mounted vertically in the

center-placed hole of the base (step 1). Then the Jumping Jack tiger/crocodile is attached to the stick due to the Velcro parts on both objects (step 2). Finally, the string on the Jumping Jack is pulled causing it to jump (step 3).

The Spinner consisted of a blue square-shaped 24.5×24.5 cm base with two 15 cm tall blue poles (of which one was firm, whereas the other was hinged) and a 38 cm long green stick with a red center located, horizontally mounted "wheel" (diameter = 11 cm). The analog version of the Spinner had a yellow 6-edged base (maximum distance between edges = 24 cm); the two yellow poles were 15 cm tall, and the 38 cm long stick was black and had a blue "spinner" (disk diameter = 11 cm). The three-step sequence to be remembered for the Spinner was the following: first, you raise the hinged pole to an upright position (step 1). Then you place the stick on the poles as a cross bar (step 2). Finally, you turn the spinner (step 3).

The Shaker consisted of two half ball shells (one in pale red and one in light blue; diameter = 9 cm) and a white ball (4 cm in diameter). The analog version of the Shaker consisted of two half ball shells (one in pink and one in yellow; diameter = 9 cm) and a yellow ball (diameter = 4 cm). The three-step sequence to be remembered for the Shaker was as follows: first you place the ball in one of the half ball shells (step 1). Then you cover up the ball by mounting the remaining half ball shell to the other half ball shell (step 2). Finally, you shake the ball with both hands (step 3).

The Windmill consisted of a square-shaped green 17.5×17.5 cm base with a center located hole covered by a wooden plate that could be moved to the side in order to uncover the mounting hole; a yellow 22 cm tall windmill with three wings (14 cm) colored in white, green, and red. The analog version of the Windmill consisted of a 6-edged pink base (maximum distance between edges = 20 cm), with a horizontally mounted blue disk (11 cm in diameter) covering the mounting hole for the windmill; a red 22 cm tall windmill with three wings (14 cm) in grey, black, and blue. The three-step sequence to be remembered for the Windmill was the following: first you uncover the center located hole by sliding the plate to the side (step 1) (or in the analog version by turning the wheel). Then you mount the windmill into the hole of the base (step 2). Finally, you turn the wings of the windmill (step 3).

In order to be able to control for potential differences in language abilities, all parents were instructed to fill out and hand in a standardized Danish version of the MacArthur-Bates Communicative Development Inventory: Words and Sentences (CDI), when coming back for the second visit.

2.3. Procedure. Each infant visited the lab twice. On the first visit, the infant participated in a Baseline measure, a *Demonstration* of the to-be-remembered three-step sequence by the experimenter, and an Encoding test for each of the five props. The second visit, took place two weeks ($\pm$ two days) later. At the second visit, the infants' ability to reproduce the to-be-remembered events was assessed in a Delayed test after the 2-week retention period. The order of the props was counterbalanced across infants and conditions. For the sake of simplicity, we used four fixed orders across infants

and conditions: (1) Gong, Shaker, Jumping Jack, Windmill, and Spinner, (2) Windmill, Spinner, Gong, Jumping Jack, and Shaker, (3) Spinner, Windmill, Jumping Jack, and Gong, and (4) Shaker, Jumping Jack, Gong, Spinner, and Windmill. The order used was the same at both visits.

Two questions were addressed: (1) whether infants were able to generalize from the target props to the analog props and (2) whether narrative support at both baseline/demonstration and retrieval would facilitate memory.

2.3.1. Target versus Analog. At the first visit, all props belonged to the group of target props. At the second visit, two (or three) of the five props were again the target props used at the first visit, whereas the remaining three (or two) props were analog versions of the props. Which props that were target and which props that were analog on the second visit were also counterbalanced across infants and conditions. We used eight fixed combinations of two or three props to be analog at the second visit.

2.3.2. Type of Narration Accompanying the Demonstration. The infants were randomly assigned to one of four conditions: Full-Full, Full-Empty, Empty-Full, and Empty-Empty. The first adjective before the dash refers to the first visit, whereas the second adjective refers to the second visit. This means that, for instance, infants in the Full-Empty condition received a “full” narrative description at the baseline and demonstration at the first encounter, whereas the narration was “empty” at the second encounter (see below for a detailed account of the full and empty narratives, resp.).

2.3.3. Warm-Up Session. A warm-up session preceded the test at the first visit. The warm-up session served the purpose of engaging the infants in the following test. In the warm-up session, the infant was encouraged to imitate a simple sequence in which the experimenter “drank tea” with a toy cup or “tucked in the baby” with small models. When the infant was engaged in the game, the experimenter moved on to the test.

2.3.4. Structure of the Test. At the first visit, the test consisted of three parts for each prop: Baseline, Demonstration, and Encoding test. All three sequences, baseline, demonstration, and encoding were carried out consecutively for each prop. The second visit consisted only of the Delayed test. Five different assistants tested the infants, but each infant was always tested by the same assistant at the two encounters. The parents were sitting next to and slightly behind the infants during all sequences and were instructed not to interfere with the test.

2.3.5. Baseline. The infants were given all the items related to a prop on a tray accompanied with one of two different narrations: empty or full. The empty narration was as follows: “You can play with all these things. Can you show me how you play with all these things?” Whereas the empty narrations were the same across props, the full narrations differed across props, simply because the props differed. In the following, the

full narration for the Gong is given as an example: “You can use all these things to make a Gong. Can you show me how you make a Gong with all these things?” The trial lasted for 90 seconds (timed) or until the infant became engaged in nontask-oriented behavior (e.g., keep pushing the tray away or throwing items on the floor). Any potentially task-relevant behavior from the infant—including those not part of the to-be-remembered sequence—was praised by the experimenter.

2.3.6. Demonstration. The items involved for a given prop were placed next to each other in a fixed position on the tray. Each demonstration was always presented twice. Equivalent to the baseline, the demonstration for each prop came in two versions: empty or full. Whereas the motor behavior of the experimenter was completely the same across conditions, the conditions differed with regards to the content of the narration accompanying the demonstration from the experimenter. In the empty demonstration, the experimenter only gave “empty” narration (i.e., nonspecific concerning nouns, verbs, and propositions): “You can play with all these things. See how I play with all these things! I do like this! [step 1] Like this! [step 2]. And look! [step 3]. That’s how I play with all these things.” In the Full narration demonstration, on the contrary, all the nouns, verbs, and propositions related to each step were specified. Thus, whereas the empty presentations were the same across props, the full presentations differed across props, simply because the props differed. In the following, the full narration for the Gong is given as an example: “You can use all these things to make a Gong. Look how I use all these things to make a Gong! You put down the bar! [step 1] You hang up the plate! [step 2] And hit the plate! [step 3] That’s how I make a Gong with all these things!” After the first demonstration, the items were again put back to their initial position and the second demonstration began.

2.3.7. Encoding Test. The encoding test immediately followed each demonstration. The items involved in the given prop were again arranged next to each other on the tray in a fixed order and then the tray was pushed across the table to the infant accompanied with a brief narration. Again the content was different in the empty and the full conditions. In the empty condition, the experimenter said: “Can you show me how you play with all these things, just like I did?” In the Full condition the experimenter said: “Can you show me how you make a Gong with all these things, just like I did?” Note that our version of the empty narration ended with the phrase “[...] just like I did?” Thus, in this respect, our empty narration condition may implicitly have been more directive than some of the empty narrations used in similar experiments. In this respect, we may have reduced the difference between the full narration and the empty narration. The advantage was that by making the two instructions highly similar with regard to directivity and word length, we were allowed to specifically address the potential impact of using the unique nouns, verbs, and propositions related to each action step for each prop.

The trial lasted for 90 seconds (timed) or until the infant became engaged in nontask-oriented behavior. Any potentially task-relevant behavior from the infant, including those

not part of the to-be-remembered sequence, was praised by the experimenter.

2.3.8. Delayed Test (at Second Encounter). The procedure here was completely identical to the procedure for the encoding test but depended on the specific condition the infant was allocated to.

2.4. Scoring and Data Reduction. We scored the infants' behavior on the Baseline, Encoding test, and the Delayed test. Two aspects were scored: (a) the number of *steps* (re)produced and (b) the number of correctly ordered *pairs* of steps (re)produced.

- (a) *Number of Steps.* The infants received one point for each step produced. Only one point could be given for each unique step. Since all to-be-remembered sequences were three-step sequences, each infant could receive a score in the 0–3 range for each trial.
- (b) *Number of Correctly Ordered Pairs of Steps.* Here we followed the scoring strategy employed by Bauer [29]: the infants received one point for each correctly ordered pair of steps. Each infant could receive a score in the 0–2 range (if all pairs of steps were reproduced in the correct order, the infant would receive two points: one point for the step-1-step-2 order and one point for the step-2-step-3 order; alternatively, the infant could receive a single point for the combination step-1-step-3, if step 2 was left out or was produced out of sequence). Only the first occurrence of each step was included. If, for instance, the infant reproduced all three steps involved in making the Gong, but in the order 3-1-2-3: *And hit the plate!* [step 3] *You put down the bar!* [step 1] *You hang up the plate!* [step 2], *And hit the plate!* [step 3] then the infant would receive a score of 3 for producing all the target steps, but only score 1 for the correctly ordered pair step-1-step-2. The order step-2-step-3 was not credited because step 3 had already occurred. This way of calculating the number of correctly ordered pairs reduces the likelihood that the infants would produce correctly order pairs by trial and error or by chance. Note that the two scores, the number of steps and the number of correctly ordered steps, are not independent of one another [29].

All trials were scored by a single primary scorer based on substantial training with a scoring manual. Due to the nature of the design, it was impossible to blind the scorer from the hypothesis of the present study. In order to objectify the scoring process, we therefore chose to have a second independent scorer rescore all trials (instead of just 20–25% as is usually done) based on the same scoring manual and after having received substantial training. The interrater agreement was 97.5%; Pearson's, $r = .99$.

Since the scoring of the different assistants did not differ, mean scores for number of steps and number of correctly ordered pairs for the infants were averaged and calculated

across assistants. Subsequent analyses are based on these means.

3. Results

The means and standard deviations for the number of correct steps and the number of correctly ordered pairs of steps across props for Baseline, Encoding test, and Delayed test, respectively, and for all conditions are presented in Table 1.

Initially, we analyzed whether the design worked as planned; that is, whether the infants as a whole group actually remembered the to-be-remembered sequences. We first looked at the number of *steps* to be remembered. A one-way repeated-measures ANOVA with Measure (Baseline versus Encoding test versus Delayed test) as within-subjects factor and with the number of *steps* produced as dependent variable yielded a strong significant effect, $F(2, 142) = 317.03$, $P < .0001$, $\eta_p^2 = .82$. Subsequent pairwise post hoc tests using the Bonferroni correction revealed that the infants overall clearly produced a significantly ($P < .0001$) larger number of correct steps at the Encoding test ($M_{\text{Step-Encoding}} = 2.59$, $SD = 0.42$) immediately following the demonstration, compared to Baseline ($M_{\text{Step-Base}} = 0.91$, $SD = 0.39$). The infants as a group were also clearly able to produce a significantly ($P < .0001$) larger number of to-be-remembered steps at the Delayed test across the two-week retention interval ($M_{\text{Step-Delay}} = 1.96$, $SD = 0.55$) relative to Baseline ($M_{\text{Step-Base}} = 0.91$, $SD = 0.39$). Meanwhile, the infants as a group also displayed evidence of reliable forgetting ($P < .0001$) over the two-week retention interval ($M_{\text{Step-Delay}} = 1.96$, $SD = 0.55$; $M_{\text{Step-Encoding}} = 2.59$, $SD = 0.42$).

The same pattern in results was displayed when analyzing the number of correctly ordered *pairs* remembered. A one-way repeated-measures ANOVA with Measure (Baseline versus Encoding test versus Delayed test) as within-subjects factor and with the number of correctly ordered *pairs* produced as a dependent variable yielded a strong significant effect, $F(2, 142) = 292.78$, $P < .0001$, $\eta_p^2 = .81$. Subsequent pairwise post hoc tests using the Bonferroni procedure revealed that the infants as a group clearly produced significantly ($P < .0001$) more ordered pairs of steps at the Encoding test ($M_{\text{Pair-Encoding}} = 1.55$, $SD = 0.39$) compared to Baseline ($M_{\text{Pair-Base}} = 0.16$, $SD = 0.17$). The infants as a group were also able to produce a significantly ($P < .0001$) larger number of correctly ordered pairs of steps over the two-week retention period at the Delayed test ($M_{\text{Pair-Delay}} = 0.99$, $SD = 0.46$) relative to Baseline ($M_{\text{Pair-Base}} = 0.16$, $SD = 0.17$). However, the infants also provided evidence of forgetting ($P < .0001$) with regard to the number of correctly ordered pairs across the two-week retention interval, ($M_{\text{Pair-Delay}} = 0.99$, $SD = 0.46$; $M_{\text{Pair-Encoding}} = 1.55$, $SD = 0.39$).

Thus, the infants demonstrated that they had learned the sequences and could remember them across the two-week retention interval but also showed evidence of forgetting between the Encoding test and the Delayed test.

Because the two experimental manipulations were conducted at the same time on the same subjects, we analyzed their possible effects by means of two ANOVAs involving

TABLE 1: The means and standard deviations for the number of correct steps and the number of correctly ordered pairs of steps across props for Baseline, Encoding test, and Delayed test, respectively, for all conditions.

Condition	<i>n</i>	Test											
		Baseline				Encoding test				Delayed test			
		Correct steps		Correct pairs		Correct steps		Correct pairs		Correct steps		Correct pairs	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
All	72	0.91	0.39	0.16	0.17	2.59	0.42	1.55	0.39	1.96	0.55	0.99	0.46
Target	72	0.91	0.39	0.16	0.17	2.59	0.42	1.55	0.39	2.03	0.76	1.04	0.65
Analogue	72	—	—	—	—	—	—	—	—	1.87	0.66	0.90	0.55
Full-Full, total	18	0.79	0.45	0.16	0.18	2.50	0.44	1.44	0.43	1.59	0.44	0.71	0.36
Full-Full, target	18	—	—	—	—	—	—	—	—	1.72	0.66	0.77	0.56
Full-Full, analog	18	—	—	—	—	—	—	—	—	1.44	0.63	0.64	0.50
Full-Empty, total	18	0.96	0.43	0.17	0.18	2.56	0.48	1.50	0.40	2.00	0.55	0.99	0.49
Full-Empty, target	18	—	—	—	—	—	—	—	—	1.99	0.80	0.99	0.65
Full-Empty, analog	18	—	—	—	—	—	—	—	—	1.94	0.58	0.94	0.54
Empty-Full, total	18	1.06	0.36	0.17	0.17	2.78	0.25	1.72	0.28	2.22	0.47	1.16	0.41
Empty-Full, target	18	—	—	—	—	—	—	—	—	2.34	0.61	1.29	0.64
Empty-Full, analog	18	—	—	—	—	—	—	—	—	2.12	0.60	1.03	0.44
Empty-Empty, total	18	0.86	0.26	0.14	0.17	2.54	0.45	1.52	0.42	2.03	0.56	1.07	0.47
Empty-Empty, target	18	—	—	—	—	—	—	—	—	2.06	0.88	1.13	0.69
Empty-Empty, analog	18	—	—	—	—	—	—	—	—	1.98	0.66	0.99	0.64

both experimental conditions on (1) mean number of *steps* at Delayed test and (2) mean number of correctly ordered *pairs* at Delayed test. In both cases, the measure of infants' receptive language abilities, CDI-R, was used as covariate.

First, we conducted a 4×2 mixed-model ANOVA with *Condition* (Full-Full versus Full-Empty versus Empty-Full versus Empty-Empty) as between-subjects factor and *Generalization* (Target props versus Analog props) as within-subjects factor and with CDI-R added as a covariate. The dependent variable was the number of correctly produced *steps* at the Delayed test. This analysis yielded a single main effect of *Condition*, $F(3, 68) = 5.14$, $P < .005$, $\eta_p^2 = .19$. Somewhat surprisingly, the infants in the Full-Full condition produced the lowest number of correct steps, whereas the infants in the Empty-Full condition were most successful ($M_{\text{Full-Full-Step-Delayed}} = 1.59$, $SD = 0.44$; $M_{\text{Full-Empty-Step-Delayed}} = 2.00$, $SD = 0.55$; $M_{\text{Empty-Full-Step-Delayed}} = 2.22$, $SD = 0.47$; $M_{\text{Empty-Empty-Step-Delayed}} = 2.03$, $SD = 0.56$). Pairwise post hoc tests using the Bonferroni correction revealed that the Full-Full condition ($M_{\text{Full-Full-Step-Delayed}} = 1.59$, $SD = 0.44$) was significantly different ($P < .005$) from the Empty-Full condition ($M_{\text{Empty-Full-Step-Delayed}} = 2.22$, $SD = 0.47$). No other main effect or interactions were obtained (all P 's $> .05$).

Note, that the nonsignificant main effect of the within-subjects factor *Generalization* ($F[1, 67] = .98$, $P = .33$, $\eta_p^2 = .01$) indicated that the infants did not produce reliably fewer correct steps with the analog props at the Delayed test ($M_{\text{Analog-Step-Delayed}} = 1.87$, $SD = 0.66$) compared to the number of correct steps with the target props that ($M_{\text{Target-Step-Delayed}} = 2.03$, $SD = 0.76$) they were presented with during the demonstration. Thus, the infants were indeed able to generalize from the target props to the analog props. The infants' successful generalization to the analog

props was further underscored by the results from a paired samples *t*-test revealing that the infants produced reliably more correct steps with analog props at the Delayed test ($M_{\text{Analog-Step-Delayed}} = 1.87$, $SD = 0.66$) compared to Baseline ($M_{\text{Step-Base}} = 0.91$, $SD = 0.39$), $t(71) = -11.06$, $P < .0001$, $r = .80$ (2-tailed). Thus, the infants were indeed able to generalize from the target props to the analog props across a two-week retention interval (see Figure 2).

Note further that the nonsignificant interaction ($F(3, 67) < .53$) between *Condition* and *Generalization* indicated that the different kinds of narrative support did not interact systematically with the infants' ability to generalize from the target props to the analog props with regard to the number of correct steps produced at the Delayed test. Finally, the fact that CDI-R measures were not involved in any significant interactions revealed that the receptive vocabulary of the infants had no systematic effect on the results.

Second, we repeated the analysis but this time with regard to the number of correctly ordered *pairs* produced at the Delayed test. This was done by means of a 4×2 mixed-model ANOVA with *Condition* (Full-Full versus Full-Empty versus Empty-Full versus Empty-Empty) as between-subjects factor and *Generalization* (Target props versus Analog props) as within-subjects factor, and again with CDI-R as covariate, but this time with the number of correctly ordered *pairs* at the Delayed test as dependent variable. The analysis revealed a single main effect of *Condition*, $F(3, 67) = 3.57$, $P < .05$, $\eta_p^2 = .14$. Again, somewhat surprisingly, the infants in the Full-Full condition produced the lowest number of correctly ordered pairs while the infants in the Empty-Full condition produced the highest number of correctly ordered pairs ($M_{\text{Full-Full-Pair-Delayed}} = 0.71$, $SD = 0.36$; $M_{\text{Full-Empty-Pair-Delayed}} = 0.99$, $SD = 0.49$; $M_{\text{Empty-Full-Pair-Delayed}} = 1.16$, $SD = 0.41$;

FIGURE 2: Graphical presentation of the mean number of correct steps produced at Baseline, Encoding test, and Delayed test for both Target props and Analogue props. Error bars ± 1 SE, * $P < .0001$.

$M_{\text{Empty-Empty-Pair-Delayed}} = 1.07$, $SD = 0.47$). Pairwise post hoc tests using the Bonferroni correction revealed that the Full-Full condition ($M_{\text{Full-Full-Pair-Delayed}} = 0.71$, $SD = 0.36$) was significantly different ($P < .05$) from Empty-Full condition ($M_{\text{Empty-Full-Pair-Delayed}} = 1.16$, $SD = 0.41$).

In close resemblance to the results obtained with the number of correct steps produced, the present analysis also yielded a nonsignificant main effect of the within-subjects factor *Generalization* ($F(1, 67) = 1.89$, $P = .17$, $\eta_p^2 = .03$), but this time with regard to correctly ordered pairs ($M_{\text{Analog-Pair-Delay}} = 0.90$, $SD = 0.55$; $M_{\text{Target-Pair-Delay}} = 1.04$, $SD = 0.65$). Note again that the nonsignificant result indicated that the infants were indeed able to *generalize* from the target props to the analog props at the Delayed test across a two-week retention interval (see Figure 2). Again, this pattern of results was supported from a paired samples *t*-test revealing that the infants produced significantly more correctly ordered pairs of steps with the analog props at the Delayed test ($M_{\text{Analog-Pair-Delayed}} = 0.90$, $SD = 0.55$) compared to Baseline ($M_{\text{Pair-Base}} = 0.16$, $SD = 0.17$), $t(71) = -10.86$, $P < .0001$, $r = .79$ (see Figure 3).

The nonsignificant interaction ($F(3, 67) < .36$) between *Condition* and *Generalization* indicated that the different degree of narrative support did not interact systematically with the infants ability to generalize from the target props to the analog props with regard to the number of correctly ordered pairs at the Delayed test.

However, in contrast to the equivalent analysis with regard to steps, the analysis on correctly ordered pairs did yield a significant interaction between *Generalization* and CDI-R, $F(1, 67) = 4.30$, $P < .05$, $\eta_p^2 = .06$. In order to qualify the interaction two additional correlational analyses

FIGURE 3: Graphical presentation of the mean number of correctly ordered pairs produced at Baseline, Encoding test, and Delayed test for both Target props and Analogue props. Error bars ± 1 SE, * $P < .0001$.

were conducted. These revealed that while CDI-R scores correlated positively with the number of correctly ordered pairs produced with the *target props* (Pearson's, $r = .17$, $P = .16$), CDI-R scores correlated negatively with the number of correctly ordered pairs produced with the analog props (Pearson's, $r = -.13$, $P = .27$). Both correlations are relatively weak and nonsignificant, and since no obvious explanation can be provided, no further analysis was conducted.

4. Discussion

The present study is the first to provide clear evidence that 18-month-old infants are indeed capable of generalizing to differently looking but functionally equivalent props across a retention interval of two full weeks with regard to both the number of correctly produced steps as well as correctly ordered pairs of steps.

However, contrary to our hypothesis, narrative support did not enhance the memory abilities of the infants. First, there was no interaction between narrative support and the infants' ability to generalize. Second, and contrary to the expectations based on the relatively few existing studies, the infants in the FF condition receiving narrative support at both Demonstration and Delayed test actually fared worse than their equally aged peers in the EF condition. How do we explain these results? In the following each of the two main results will be discussed in turn.

Why did we obtain successful generalizing in 18-month-old infants across a retention interval of two full weeks? First, by means of employing the elicited imitation paradigm, the infants in the present study were given the opportunity to manipulate the objects at the first visit to the lab, that is,

both at baseline and during the encoding test. In this respect, the present study resembles the procedure employed in the Bauer and Dow [19] study, where successful generalization to differently looking but functionally equivalent props with both 16- and 20-month-old infants was obtained across a retention interval of one week. In comparison, the 18-month-old infants in the Herbert and Hayne [18] study failed to generalize across a 24 h retention interval, whereas their 24- and 30-month-old peers succeeded. However, the Herbert and Hayne [18] study was an “observation only” study where the experimental group was tested by means of the deferred imitation paradigm meaning that the infants were given no opportunity to manipulate the objects before the delayed test. Given the fact that practice does indeed facilitate memory performance in imitation paradigms [30, 31], we believe that this difference may be at least partly responsible for the positive results obtained in the present study across a retention interval of two weeks.

Second, although our analog props differed from the target props with respect to shape and color, the differences could have been of a larger magnitude. In principle, a change in color can be anything from, for instance, a slight, yet detectable, change within the broad family of the color “red,” to a radical change from “white” to “black.” Similarly, and probably more influential (cf. [32]), a shape change can in principle be anything from adding (or removing) a single, tiny edge, to radical changes like changing a cylinder-shaped stick serving the function as a handle to a complex uneven surfaced object where only the functional properties are preserved. Consider, for instance, the variety of shapes that everyday objects like glasses or candlesticks come in while still sharing a functional core within their respective categories. Recall that the 18-month-old infants in Herbert and Hayne’s [18] study were unable to generalize to differently looking but functionally equivalent props after a 24 h delay, whereas their 24- and 30-months-old peers succeeded. The failure of the 18-month-old infants in Herbert and Hayne’s [18] study may be due to the rather substantial shape changes between the target and the analog props. The handle of the rattle used in the Herbert and Hayne [18, page 475] study, for example, was quite different with regards to shape from target prop to analog prop: one of the handles was just a cylinder shaped stick, whereas the other handle consisted of two-connected sticks mounted in parallel. Relative to such shape changes, the shape changes involved in our study were minor (e.g., from a square shaped base to a 6-edged shaped base). The shape changes employed in the present study simply had stronger resemblance to the shape changes used in Bauer and Dow’s [19, Appendix] study, which also showed successful generalization across a one-week retention interval in both 16- and 20-month-old infants. To summarize, we believe that the combination of giving the infants the opportunity to manipulate the props prior to the delayed test and the relatively minor featural changes employed may explain why the infants in the present study were able to generalize to differently looking but functionally equivalent props across a retention interval of two weeks.

Why did narrative support fail to facilitate generalization in the present study? As outlined in the introduction,

narrative support has been shown to have a positive effect on infants and children’s memory abilities [24, 26, 27]. However, most of the evidence has been obtained with children and not with infants. Furthermore, at least two of the few existing studies on generalization with infants revealed that 18-month-old infants did not benefit from being provided with narrative support [18, Exp. 2], [28]. Moreover, the results from Hayne and Herbert [27] revealed that although infants receiving full narration overall fared significantly better than infants receiving empty narration, narration had no effect on the encoding sequences but only influenced the infants’ performance on the delayed test. Finally, infants in the EF condition were more successful than the infants in the FE condition [18]—just as it was the case in the present study. To summarize, although this is not a general pattern, evidence indicating that narrative support may have no facilitating effect on infants does exist. The results from the present study are in accordance with this minority.

We consider two possible explanations for the results obtained with regard to narrative support having no effect on the infants’ ability to generalize: first, our operationalization of the Empty narration was somewhat more directive than is usually seen in similar studies. As specified in Section 2, our Delayed tests—regardless of whether it was an empty or a full narration—ended with the implicitly directive phrase “[...] *just like I did.*” Consequently, the difference between the empty and the full accompanying narrative was reduced to a difference with regard to the use of prop-specific nouns, verbs, and propositions, whereas the strength of the intentional request was kept very similar between the two kinds of narrative support. Thus, it may be that the core of the overall intentional message involved in the narrative support (i.e., “[...] *just like I did.*”) is more important for an 18-month-old infant in order to reproduce the to-be-remembered action sequence than specifying the nouns, verbs, and propositions related to the fairly novel props. This tentative interpretation is in accordance with the recent results obtained by Simcock and colleagues [16] cited in the introduction where infants receiving explicit instructions “Show me how to make a rattle” fared significantly better than infants who received nondirective prompts “Show me how to make something” [16]. The term “something” may not be understood by the infant as a “place holder” for a given unspecified object but rather as an invitation to do “anything you like,” which lacks directivity. The tentative interpretation offered here is also in accordance with the recent trend that within infants’ social cognitive development where it has become abundantly clear that infants in their second year of life are indeed competent “intention readers” (for reviews, see e.g., [33, 34]). However, the present interpretation is indeed suggestive and further experiments are clearly needed in order to reach a firm conclusion regarding the relative impact of uses of, on the one hand, directive phrases and, on the other hand, specific prop-related nouns, verbs, and propositions.

Second and closely related to the tentative explanation outlines above, it might be that until the infants’ language abilities have matured, providing explicit names of the props may not always be the most effective kind of narrative support and may even at times interfere with the infants’

attempts to encode or retrieve a to-be-remembered sequence (cf. [28]).

In a broader perspective, studies on infants' and children's ability to generalize knowledge from one set of objects to another may be used to qualify theories on situated learning. Theories on situated learning have argued that most cognitive theories of learning tend to disregard the importance of the context in which a given skill is acquired and that the difficulties that human beings face when attempting to employ a given skill in a different context than the one where the skill was originally learned have been underestimated (e.g., [35]). This may be true. However, when comparing the achievements of human beings to any other species with regard to the ability to *abstract* knowledge from one domain and to employ this knowledge in a different domain, the differences can in our view not be overestimated. The ability to abstract core functional aspects of a given object may even be considered a crucial feature of what makes human beings unique [1, 36]. Studies from developmental psychology demonstrate that the ability to generalize functional attributes across featural dissimilarities and retention intervals is an early developmental achievement [14, 18, 19, 26] underscoring the enormous flexibility that we as human beings seem to be endowed with.

In conclusion, the present study provides the first evidence that 18-month-old infants are capable of generalizing to differently looking but functionally equivalent props across a retention interval of two full weeks. In the present study, narrative support had no facilitating effect, and further research is needed in order to clarify when and under what specific circumstances narrative support may facilitate infant memory abilities.

Acknowledgments

This study was supported by the Danish National research foundation (Grant no. DNRF93) and MINDLab. The authors would like to thank Jonna Jelsbak Dahl, Inger Birchall Nielsen, Ida Marie Lautrup, Line R. Kristiansen, Bettina Hald, and Trine Sonne for their help collecting and scoring the data. They would also like to thank Patricia Bauer for many constructive suggestions all through the study as well as for giving them access to details regarding the props. They are also grateful for the constructive comments on the paper from the three anonymous reviewers and Trine Sonne. Finally, they would like to thank all the infants who participated in the study as well as their parents for allowing them to do so.

References

- [1] P. Krøjgaard, "The human ability to single out and track specific objects through space and time: origin and application," in *Human Characteristics*, H. Høgh-Olesen, P. Bertelsen, and J. Tønnesvang, Eds., pp. 89–116, Cambridge Scholars Publishing, 2009.
- [2] D. Gentner and K. J. Holyoak, "Reasoning and learning by analogy," *The American psychologist*, vol. 52, no. 1, pp. 32–34, 1997.
- [3] P. J. Bauer, *Remembering the Times of Our Lives: Memory in Infancy and Beyond*, Lawrence Erlbaum Associates, Mahwah, NJ, USA, 2007.
- [4] H. Hayne, "Infant memory development: implications for childhood amnesia," *Developmental Review*, vol. 24, no. 1, pp. 33–73, 2004.
- [5] M. L. Howe, *The Nature of Early Memory: An Adaptive Theory of the Genesis and Development of Memory*, Oxford University Press, New York, NY, USA, 2011.
- [6] C. Rovee-Collier and K. Cuevas, "The development of infant memory," in *The Development of Memory in Infancy and Childhood*, M. L. Courage and N. Cowan, Eds., pp. 11–41, Psychology Press, New York, NY, USA, 2nd edition, 2009.
- [7] M. W. Sullivan, C. K. Rovee-Collier, and D. M. Tynes, "A conditioning analysis of infant long-term memory," *Child development*, vol. 50, no. 1, pp. 152–162, 1979.
- [8] W. L. Hill, D. Borovsky, and C. Rovee-Collier, "Continuities in infant memory development," *Developmental Psychobiology*, vol. 21, no. 1, pp. 43–62, 1988.
- [9] H. Hayne, C. Greco, L. Earley, P. Griesler, and C. Rovee-Collier, "Ontogeny of early event memory: II. Encoding and retrieval by 2- and 3-month-olds," *Infant Behavior and Development*, vol. 9, no. 4, pp. 461–472, 1986.
- [10] D. Borovsky and C. Rovee-Collier, "Contextual constraints on memory retrieval at six months," *Child development*, vol. 61, no. 5, pp. 1569–1583, 1990.
- [11] B. S. DeFrancisco and C. Rovee-Collier, "The specificity of priming effects over the first year of life," *Developmental Psychobiology*, vol. 50, no. 5, pp. 486–501, 2008.
- [12] A. N. Meltzoff, "What infant memory tells us about infantile amnesia: long-term recall and deferred imitation," *Journal of Experimental Child Psychology*, vol. 59, no. 3, pp. 497–515, 1995.
- [13] H. Hayne and G. Simcock, "Memory development in toddlers," in *The Development of Memory in Infancy and Childhood*, M. L. Courage and N. Cowan, Eds., pp. 43–68, Psychology Press, New York, NY, USA, Second edition, 2009.
- [14] S. B. Barnat, P. J. Klein, and A. N. Meltzoff, "Deferred imitation across changes in context and object: memory and generalization in 14-month-old infants," *Infant Behavior and Development*, vol. 19, no. 2, pp. 241–251, 1996.
- [15] P. J. Klein and A. N. Meltzoff, "Long-term memory, forgetting, and deferred imitation in 12-month-old infants," *Developmental Science*, vol. 2, no. 1, pp. 102–113, 1999.
- [16] G. Simcock, K. Garrity, and R. Barr, "The effect of narrative cues on infants' imitation from television and picture books," in *Child Development*, vol. 82, pp. 1607–1619, 2011.
- [17] N. Brito, R. Barr, P. McIntyre, and G. Simcock, "Long-term transfer of learning from books and video during toddlerhood," *Journal of Experimental Child Psychology*, vol. 111, pp. 108–119, 2012.
- [18] J. Herbert and H. Hayne, "Memory retrieval by 18–30-month-olds: age-related changes in representational flexibility," *Developmental Psychology*, vol. 36, no. 4, pp. 473–484, 2000.
- [19] P. J. Bauer and G. A. Dow, "Episodic memory in 16- and 20-month-old children: specifics are generalized but not forgotten," *Developmental Psychology*, vol. 30, no. 3, pp. 403–417, 1994.
- [20] F. Jack, S. MacDonald, E. Reese, and H. Hayne, "Maternal reminiscing style during early childhood predicts the age of adolescents' earliest memories," *Child Development*, vol. 80, no. 2, pp. 496–505, 2009.
- [21] R. Fivush, C. A. Haden, and E. Reese, "Elaborating on elaborations: role of maternal reminiscing style in cognitive and

- socioemotional development,” *Child Development*, vol. 77, no. 6, pp. 1568–1588, 2006.
- [22] E. Reese and R. Newcombe, “Training mothers in elaborative reminiscing enhances children’s autobiographical memory and narrative,” *Child Development*, vol. 78, no. 4, pp. 1153–1170, 2007.
 - [23] F. McGuigan and K. Salmon, “The time to talk: the influence of the timing of adult-child talk on children’s event memory,” *Child Development*, vol. 75, no. 3, pp. 669–686, 2004.
 - [24] M. Tessler and K. Nelson, “Making memories: the influence of joint encoding on later recall by young children,” *Consciousness and Cognition*, vol. 3, no. 3-4, pp. 307–326, 1994.
 - [25] P. J. Bauer, J. A. Wenner, P. L. Dropik, and S. S. Wewerka, “Parameters of remembering and forgetting in the transition from infancy to early childhood,” *Monographs of the Society for Research in Child Development*, vol. 65, no. 4, pp. 1–203, 2000.
 - [26] J. Herbert, “The effect of language cues on infants’ representational flexibility in a deferred imitation task,” *Infant Behavior and Development*, vol. 34, pp. 632–635, 2011.
 - [27] H. Hayne and J. Herbert, “Verbal cues facilitate memory retrieval during infancy,” *Journal of Experimental Child Psychology*, vol. 89, no. 2, pp. 127–139, 2004.
 - [28] E. Zack, P. Gerhardstein, A. N. Meltzoff, and R. Barr, “15-month-olds’ transfer of learning between touch screen and real-world displays: language cues and cognitive loads,” *Scandinavian Journal of Psychology*, vol. 54, pp. 20–25, 2012.
 - [29] P. J. Bauer, “Holding it all together: how enabling relations facilitate young children’s event recall,” *Cognitive Development*, vol. 7, no. 1, pp. 1–28, 1992.
 - [30] H. Hayne, R. Barr, and J. Herbert, “The effect of prior practice on memory reactivation and generalization,” *Child Development*, vol. 74, no. 6, pp. 1615–1627, 2003.
 - [31] A. E. Learmonth, R. Lamberth, and C. Rovee-Collier, “Generalization of deferred imitation during the first year of life,” *Journal of Experimental Child Psychology*, vol. 88, no. 4, pp. 297–318, 2004.
 - [32] T. Wilcox, “Object individuation: infants’ use of shape, size, pattern, and color,” *Cognition*, vol. 72, no. 2, pp. 125–166, 1999.
 - [33] M. Tomasello, M. Carpenter, J. Call, T. Behne, and H. Moll, “Understanding and sharing intentions: the origins of cultural cognition,” *Behavioral and Brain Sciences*, vol. 28, no. 5, pp. 675–691, 2005.
 - [34] G. Csibra and G. Gergely, ““Obsessed with goals”: functions and mechanisms of teleological interpretation of actions in humans,” *Acta Psychologica*, vol. 124, no. 1, pp. 60–78, 2007.
 - [35] J. Lave and E. Wenger, *Situated Learning: Legitimate Peripheral Participation*, University of Cambridge Press, Cambridge, UK, 1991.
 - [36] O. S. Kingo and P. Krøjgaard, “Object function as a basic requirement for language use and tool use,” in *Human Characteristics*, H. Høgh-Olesen, J. Tønnesvang, and P. Bertelsen, Eds., pp. 65–88, Cambridge Scholars Publishing, Cambridge Scholars Publishing, UK, 2009.