Cognitive Psychology 75 (2014) 55–79

Contents lists available at ScienceDirect

Cognitive Psychology

journal homepage: www.elsevier.com/locate/cogpsych

Developmental origins of recoding and decoding in memory

Melissa M. Kibbe [a,b,*], Lisa Feigenson [b]

[a] Department of Psychological and Brain Sciences, Boston University, 64 Cummington Mall, Boston, MA 02215, USA
[b] Department of Psychological and Brain Sciences, Johns Hopkins University, 3400 N Charles St, Baltimore, MD 21218, USA

ARTICLE INFO

Article history:
Accepted 1 August 2014

Keywords:
Working memory
Long-term memory
Chunking
Development

ABSTRACT

Working memory is severely limited in both adults and children, but one way that adults can overcome this limit is through the process of recoding. Recoding happens when representations of individual items are chunked together into a higher order representation, and the chunk is assigned a label. That label can then be decoded to retrieve the individual items from long-term memory. Whereas this ability has been extensively studied in adults (as, for example, in classic studies of memory in chess), little is known about recoding's developmental origins. Here we asked whether 2- to 3-year-old children also can recode—that is, can they restructure representations of individual objects into a higher order chunk, assign this new representation a verbal label, and then later decode the label to retrieve the represented individuals from memory. In Experiments 1 and 2, we showed children identical blocks that could be connected to make tools. Children learned a novel name for a tool that could be built from two blocks, and for a tool that could be built from three blocks. Later we told children that one of the tools was hidden in a box, with no visual information provided. Children were allowed to search the box and retrieve varying numbers of blocks. Critically, the retrieved blocks were identical and unconnected, so the only way children could know whether any blocks remained was by using the verbal label to recall how many objects comprised each tool (or chunk). We found that even children who could not yet count adjusted their searching of the box depending on the label they had heard. This suggests that they had recoded representations of individual blocks into higher-order chunks, attached labels to the chunks,

---

* Corresponding author at: Department of Psychological and Brain Sciences, Boston University, 64 Cummington Mall, Boston, MA 02215, USA.

E-mail addresses: kibbe@bu.edu (M.M. Kibbe), feigenson@jhu.edu (L. Feigenson).

http://dx.doi.org/10.1016/j.cogpsych.2014.08.001

and then later decoded the labels to infer how many blocks were hidden. In Experiments 3 and 4 we asked whether recoding also can expand the number of individual objects children could remember, as in the classic studies with adults. We found that when no information was provided to support recoding, children showed the standard failure to remember more than three hidden objects at once. But when provided recoding information, children successfully represented up to five individual objects in the box, thereby overcoming typical working memory limits. These results are the first demonstration of recoding by young children; we close by discussing their implications for understanding the structure of memory throughout the lifespan.

1. Introduction

Many years ago, George Miller (1956) made an influential observation about the limits of humans' ability to remember information over short intervals. He noted that the span of immediate memory could not be characterized in terms of a discrete amount of information measurable in, say, number of bits. Rather, immediate memory appears to hold a fixed number of *chunks*, with each chunk holding an effectively limitless amount of information, thanks to the process of *recoding.* Recoding involves taking some input (e.g., a string of numbers such as 070302215911) and dividing the input into meaningful chunks (e.g., my mother's birthday, my university's zip code, the emergency telephone number). If one wants to remember such a string of numbers, and if those numbers have been linked with semantic content, one need only maintain the much shorter list of recoded units in working memory. To retrieve the individual numbers, one *decodes* those units using knowledge stored in long-term memory. The hierarchical organization of recoded information allows for the compression of that information without serious informational degradation or loss (Dirlam, 1972; Shannon, 1948).

The psychological processes of recoding and decoding have been studied in adults in now-classic experiments on memory in the game of chess (Charness, 1976; Chase & Simon, 1973; De Groot, 1965; Frey & Adesman, 1976). For example, Chase and Simon (1973) briefly presented adult observers with chess boards on which the pieces were placed either randomly or in positions that could be part of a playable chess game, and asked them to reconstruct the boards from memory. They found that when observers were shown playable boards, the accuracy with which the pieces were reconstructed was related to the observers' expertise in chess. The more advanced the chess player, the more positions they recalled. Importantly, chess experts' superior performance was not caused by greater overall working memory capacity, because when the chess pieces were placed randomly on the board there was no difference between the recall of experts and that of novices. Rather, chess experts apparently were able to *recode* the positions of the chess pieces into meaningful chunks using a coding scheme derived from their knowledge of chess. Those chunks could then be decoded to retrieve the subordinate information: the pieces and their precise positions.

Research on recoding and decoding suggests that these processes are largely automatic. Chase and Simon (1973) argued that their chess experts "saw" the board differently than did novices, perceiving the higher order relationships between the pieces instantly. Such expertise often results in differential visual processing of a display (e.g., geometric diagrams in Epelboim & Suppes, 2001; see also Chi, Glaser, & Rees, 1982, for a review of evidence from physics experts). However, despite its seeming automaticity, recoding requires at least two key steps (Bower, 1970; Ericsson & Kintsch, 1995; Simon, 1974; Wortman & Greenberg, 1971). First, representations of individual items must be *chunked* together to form a higher order unit. For example, in the case of chess, the observer must initially learn that particular configurations of pieces form a meaningful group. Critically, this new, higher order representation of the group preserves representations of each piece it contains. Second, recoding involves storing this new, higher order representation in long-term memory so that it may be accessed later on. Optionally, the recoded chunk can be given a verbal label (e.g., "Anastasia's Mate"), which can be used to reference the chunked representation. Decoding the chunk, then, requires accessing the knowledge in long-term memory, and then unpacking the representations of the individual items back into working memory. Thus, recoding and decoding require, at a minimum, the ability to segment a group of individual items from a scene, the ability to bind these individual representations into a higher-order chunk, and the ability to maintain this new representation in a durable, accessible long-term store.

Further evidence for this kind of precise recoding and decoding of information comes from other areas of expert knowledge. A famous example involves the subject S.F., who, over the course of over 200 hours in the laboratory, repeatedly was presented with lists of digits and asked to recall them (Chase & Ericsson, 1980; Ericsson, Chase, & Faloon, 1980). When S.F. began this task he could remember only about 10 digits at a time. As testing progressed, S.F. discovered that he was able to greatly increase his span by recoding groups of three and four digits into chunks, using his knowledge of, among other things, record-setting track and field times (e.g., 3492 became "3 min and 49 point 2 s, close to the world record time for running a mile") and then further grouping these recoded sequences into still larger chunks (e.g., "mile running times"). Each recoded unit then could be precisely decoded to retrieve the specific digits in the target sequence. This strategy was extremely effective; by the end of the study, S.F. was able to accurately recall lists of 80 digits.

The performance of expert chess players and of S.F. demonstrates the recoding of information in memory. But many other studies show hierarchical restructuring without recoding. Adults can use perceptual properties of stimulus arrays to chunk individual items together and recall more than they could otherwise. For example, adults can recall more list items (e.g., digits, letters, or words) if presented with pauses after every third or fourth item (Bower & Winzenz, 1969; Hitch, Burgess, Towse, & Culpin, 1996; Ryan, 1969), and can remember the spatial locations of more objects if objects are perceptually groupable by proximity (Sargent, Dopkins, Philbeck, & Chichka, 2010). Adults also are sensitive to the statistical properties of arrays, and can use this sensitivity to increase the number of items they can recall. When stimuli exhibit statistical regularities, such as when letter strings contain redundant letters (Miller, 1958) or when runs of digits steadily increase or decrease in numerosity (Mathy & Feldman, 2012), adults can recall more than when they are presented with arrays with higher entropy (i.e., with fewer statistical regularities). Further, adults also can chunk using semantic knowledge. For example, when given lists of words to remember, they often mentally reorder the lists into sequences of words with related meanings (e.g., "cat," "dog," and "pig" can be grouped together as *animals*) (Bower, Lesgold, & Tieman, 1969; Mandler, Pearlstone, & Koopmans, 1969). As more hierarchical structure is imposed (for example, as items are grouped into larger numbers of categories, and then these categories grouped into even larger, superordinate kinds), recall improves (Wortman & Greenberg, 1971).

However, although these examples show that hierarchical restructuring enhances recall, they do not offer evidence of recoding in the sense defined by Miller (1956), because these studies did not require a precisely defined decoding scheme. For example, the category *animals* may be used to group the words "cat," "dog," and "pig," but it also picks out many other individuals not included in the stimulus list ("horse," "monkey," and so on). *Animals* cannot be precisely decoded into "cat," "dog," and "pig" because the relationship of the superordinate category to the subordinate items is not defined on the basis of specific individuals. The same is true of the other chunking schemes described above—for example, *ascending numbers* picks out an infinitely large number of possible sequences. Therefore, grouping items into familiar categories, and encoding using statistical regularities like repeated digits or runs of numbers, likely facilitate recall because these act as contextual prompts, not because the categories pick out a precise set of individual items stored in memory.

This simpler process of hierarchical restructuring *without recoding* has been shown to be available early in development. Although 7- to 8-year old children fail to spontaneously group pictures into meaningful categories when asked to remember the pictures (Rosner, 1971), children as young as 4 to 6 years old can group the pictures when given scaffolding. When shown pictures of objects from the same semantic category, or when instructed to physically sort pictures into categories, children recalled more than when the pictures were presented randomly (Kobasigawa & Orr, 1973; Sodian, Schneider, & Perlmutter, 1986). As children enter the middle school years, they can organize remembered information into more sophisticated hierarchical structures, for example by grouping items into more chunks (each containing fewer individual items) when presented with lists of words, pictures (Cowan et al., 2010), or sentences (Gilchrist, Cowan, & Naveh-Benjamin, 2009).

Chunking without recoding has been revealed in even younger children by tasks that use less explicit measures. For example, in simple object search tasks, infants typically fail to remember four individual objects at once (Barner, Thalwitz, Wood, Yang, & Carey, 2007; Feigenson & Carey, 2003, 2005). However, 14-month old infants were able to use visual grouping cues such as spatial proximity to chunk four individual objects into two sets of two, and thereby to recall all four (Feigenson & Halberda, 2004; Rosenberg & Feigenson, 2013). Similarly, even younger infants (7-month-olds) in a looking time task could remember more objects when the objects could be chunked using both spatiotemporal and featural information (Moher, Tuerk, & Feigenson, 2012). Even when spatial cues do not enable objects to be immediately grouped, infants can use objects' statistical histories to support chunking. When familiarized to four distinct objects that appeared together in random pairings, and then shown all four objects hidden together, infants failed to remember the objects. But when familiarized to the same four objects co-occurring repeatedly in predictable pairs, infants used these statistical groupings to represent the objects in chunks of two, thereby successfully remembering all four (Kibbe & Feigenson, submitted for publication). Infants also can recruit conceptual knowledge stored in long-term memory to restructure information in working memory. Whereas infants failed to remember four identical objects or four perceptually distinct objects from a single semantic category (e.g., four different cats), they succeeded when the objects could be chunked on the basis of familiar categories (e.g., two cats and two cars; Feigenson & Halberda, 2008; or two sets of interacting social partners; Stahl & Feigenson, 2014).

Although such chunking has been shown to increase children's memory performance, it too lacks a key feature of recoding and decoding as defined by Miller (1956) and as studied in expert chess players and subject S.F. Once again, this missing feature is the use of a well-defined relationship between the superordinate category and its constituents. Although children and even infants remember more when items are organized into categories like *animals*, or when objects can be grouped on the basis of spatial location, they cannot use these features to later pick out all and only the members of a specific set. This is true even when children are provided verbal labels for such categories. Labels do help young children recognize higher order relations between objects in non-numerical situations. For example, 18-month-old toddlers who were taught a novel name for an object and were then asked to "find another one" successfully picked out another token of that type, but failed to do so when the object was unlabeled (Booth & Waxman, 2002; see also Christie and Gentner (2013) for evidence that labeling can facilitate analogical reasoning). And in Waxman's classic studies of the effects of labeling on categorization, hearing objects consistently labeled by a novel noun increased the likelihood that 12- to 13-month old infants would treat a sequence of different objects as forming a coherent superordinate category (Waxman & Braun, 2005; Waxman & Markow, 1995; but see Sloutsky & Robinson, 2008). Such findings suggest that, in infancy, children use linguistic input to rapidly form categories that could affect their memory for objects.

Yet whereas, for adults, the recoded unit with the label "Anastasia's Mate" allows for the retrieval of representations of a particular set of chess pieces arranged in a particular configuration, the categories *cats* or *adjacent objects* do not permit any such precise retrieval. A category like "cat," whether labeled or not, can refer to any number of objects that can be identified as cats. Therefore, neither the previously demonstrated chunking by infants, nor infants' use of category labels to form categories of objects, meet the criteria for recoding and decoding as outlined by Miller (1956) and as studied in the expert memory literature (e.g., Chase & Simon, 1973).

There is, however, evidence that young children can perform at least one operation that is related to recoding. When children learn to count, they acquire a means of recoding representations of multiple individual objects into a single cardinal representation that specifies a precise number of individuals. The verbal labels that refer to cardinal values are stored in long-term memory, and can aid in the maintenance and retrieval of information. For example, if a set is recoded as containing "three," then a set of the remembered equivalent cardinality can later be constructed. These representations are numerically precise, in that they pick out an exact quantity – no more and no fewer than the specified number. However, this type of numerical recoding is not available until surprisingly late in childhood.

Children begin to learn the meanings of the words in their verbal count list between the ages of two and three years, with mastery usually attained around age four or even older (Gelman, 1993; Le Corre & Carey, 2007; Wynn, 1992). Furthermore, counting does not allow children to retrieve the identities of the individuals in a given set (unless the cardinality is combined with a noun, such as "four cats"); "four" can refer to any set of four individuals.

Thus, although even quite young children are capable of creating hierarchically organized memory representations, it remains unknown when in development they also can assign verbal labels to their chunked representations and use these labels to later retrieve the specific individuals in the remembered chunk—i.e., when they can recode and decode chunks in memory. To address this question, in the present studies we investigated the recoding abilities of toddlers who could not yet proficiently count. Specifically, we asked whether toddlers could bind representations of identical objects into chunks, assign verbal labels to the chunks, and finally use these labels alone, in the absence of any visual information, to retrieve representations of the individuals comprising the chunks.

We used a manual search task (Feigenson & Carey, 2003, 2005) in which toddlers had to use labels to infer how many individual objects were hidden inside a box. In a similar paradigm, Xu, Cote, and Baker (2005) showed that infants could use labels to decide how many hidden objects to search for—but in their case, no higher order representations (no chunks) were required, and each label corresponded to just one object. Here, to investigate toddlers' ability to learn labels for chunks, we first showed children identical objects that could be stuck together to form two "tools." Each tool was made from a unique number of blocks (either two or three), was demonstrated to have a unique novel function, and was given a novel verbal label. In Experiments 1 and 2, we asked whether, following this brief training experience, toddlers could decode these labels to infer how many objects were hidden in the box in the absence of visual information. In Experiments 3 and 4, we asked whether these processes of recoding and decoding also serve to *expand* children's memory (as in the case of expert chess players and S.F.), allowing them to remember more individual objects than they otherwise could.

2. Experiment 1

In Experiment 1 we asked whether toddlers have two key abilities required for recoding: first, forming a higher order chunk out of representations of individuals and attaching a label to the new chunk so that it may be later accessed from long term memory, and second, decoding the label to retrieve the individual items in the chunk. We taught toddlers that identical magnetic blocks could be stuck together to make "tools." One tool was made of two blocks and the other was made of three blocks; each tool had a novel function and each was given a novel name (for evidence that even very young infants are sensitive to objects' functions, see Wilcox & Chapa, 2004). We reasoned that assembling the blocks into tools with distinct functions would help to highlight the different numbers of blocks in each (because the 2-block tool could not perform the function of the 3-block tool, and vice versa). In the second phase of the experiment, we asked whether toddlers could use these novel labels alone (in the absence of any relevant visual information) to infer the number of blocks hidden in an opaque box. Using the labels in this way would demonstrate recoding—the binding of representations of individual items into a higher order set, the storage of the label in memory, and the ability to use the label alone to reconstruct the number of constituent objects present.

Because our aim was to discover whether toddlers can recode and decode in the sense defined by Miller, it was important to ensure that they could not simply count to solve the task. For example, if children simply mapped the tools' novel labels onto the already-known number words "two" or "three," this would not constitute evidence of recoding and decoding. Therefore, we also assessed children's counting ability using the "What's On This Card" task (Gelman, 1993; Le Corre & Carey, 2007) so that we could focus on children who could not yet count proficiently.

2.1. Method

2.1.1. Participants

Twenty-four healthy, full-term children participated (mean age: 34 months 6 days, SD: 3 months 24 days; range 24 months 17 days to 38 months, 5 days; 13 girls). Three children were identified by their parents as black, 19 as white (of whom two were identified as Hispanic or Latino), one as white and Asian, and one as "other" (Hispanic or Latino). Of the 92% of parents who self-reported their educational history, 95% had completed a college degree or higher, 3% had completed some college, and 3% had completed high school. Twelve additional children were excluded for refusal to participate in the task (4), failure to reach into the box on at least one critical test trial (3), or experimenter error (5). Children were recruited through phone lists and mailings, and received a small gift for participating.

2.1.2. Materials

The stimuli were five identical wooden blocks (3 cm × 3 cm × 3 cm) painted blue, with a hidden magnet embedded in each side so that the blocks could be connected into "tools." The functions of these tools were demonstrated using two black foam-core tunnels attached to rectangular platforms (18.5 × 14.5 cm on their base, with 6.5 × 6.5 cm walls and ceiling). In front of each tunnel was a gold sticker (5 cm diameter). Two plastic strawberries from a play food set were used to demonstrate the tools' functions (see below).

On the test trials, the blocks were hidden inside a black foam-core box (27 cm × 13 cm × 45 cm). The box had an 8 cm × 7 cm opening at the front, which was covered with yellow spandex with a slit across its width that allowed children to reach into but not see inside the box. A concealed opening at the back of the box (23 cm × 8 cm), covered with black felt, allowed the experimenter to surreptitiously reach in and withhold blocks on critical trials. Three toys were used to familiarize children with the process of retrieving objects from the box: a green plastic frog, an orange plastic tiger, and a toy shoe (all approximately 5 cm × 3 cm).

For the "What's On This Card" task used to assess children's counting ability (Gelman, 1993; Le Corre & Carey, 2007), we used four sets of six 13 cm × 21 cm white cards. Each card had images of one to six objects (apples, boats, cats, or ladybugs), printed in color in a random configuration. Each image was roughly 2.5 × 2.5 cm.

2.1.3. Procedure

Children sat in front of a low table, with the experimenter seated across from them. One video camera captured a side view of the testing session and another captured an overhead view. These images were fed into an adjacent room, where they were mixed together and digitally recorded for later coding.

2.1.3.1. Familiarization. The experiment began by familiarizing children to the box that would be used in the test trials. First, children watched as the experimenter hid the toy frog in the box; children were then allowed to reach in and retrieve it. Once children had successfully retrieved the frog, the experimenter hid the tiger and toy shoe simultaneously in the box, and children were allowed to reach in and retrieve them. After the second hidden toy had been retrieved the experimenter removed the box from the table and began the Teaching Phase.

2.1.3.2. Teaching. 2.1.3.2.1. Teaching Phase 1: Tool names. The Teaching Phase was designed to teach children the functions and names of the two tools. The experimenter began by placing the two identical foam-core tunnels on the table. She placed a toy strawberry inside one tunnel and another strawberry next to the other tunnel (Fig. 1). The left/right positions of these were counterbalanced across children.

The experimenter then demonstrated the tools' functions. First she pointed to the strawberry that was inside the tunnel and said, "See this strawberry in here? I want to put it on my plate so I can eat it, but it's stuck inside this house! What am I going to do? I know – I'm going to use my *blicker*!" She then brought out two identical blocks, which were magnetically connected in a line, and said, "See, this is my *blicker*!" To show children that the blocks were separate objects that could be conjoined, she pulled the blocks apart, waved them, and then put them back together. The experimenter then said, "It's going to *blick* the strawberry!" She used the tool to push the strawberry through the tunnel onto the gold sticker at the front of the platform, saying, "See, it's *blicking* the strawberry!" She then held up the tool, pulled apart the two blocks, conjoined them again, and said, "This is my *blicker*!" She then placed the blicker next to the platform, with its two pieces still connected (Fig. 1).

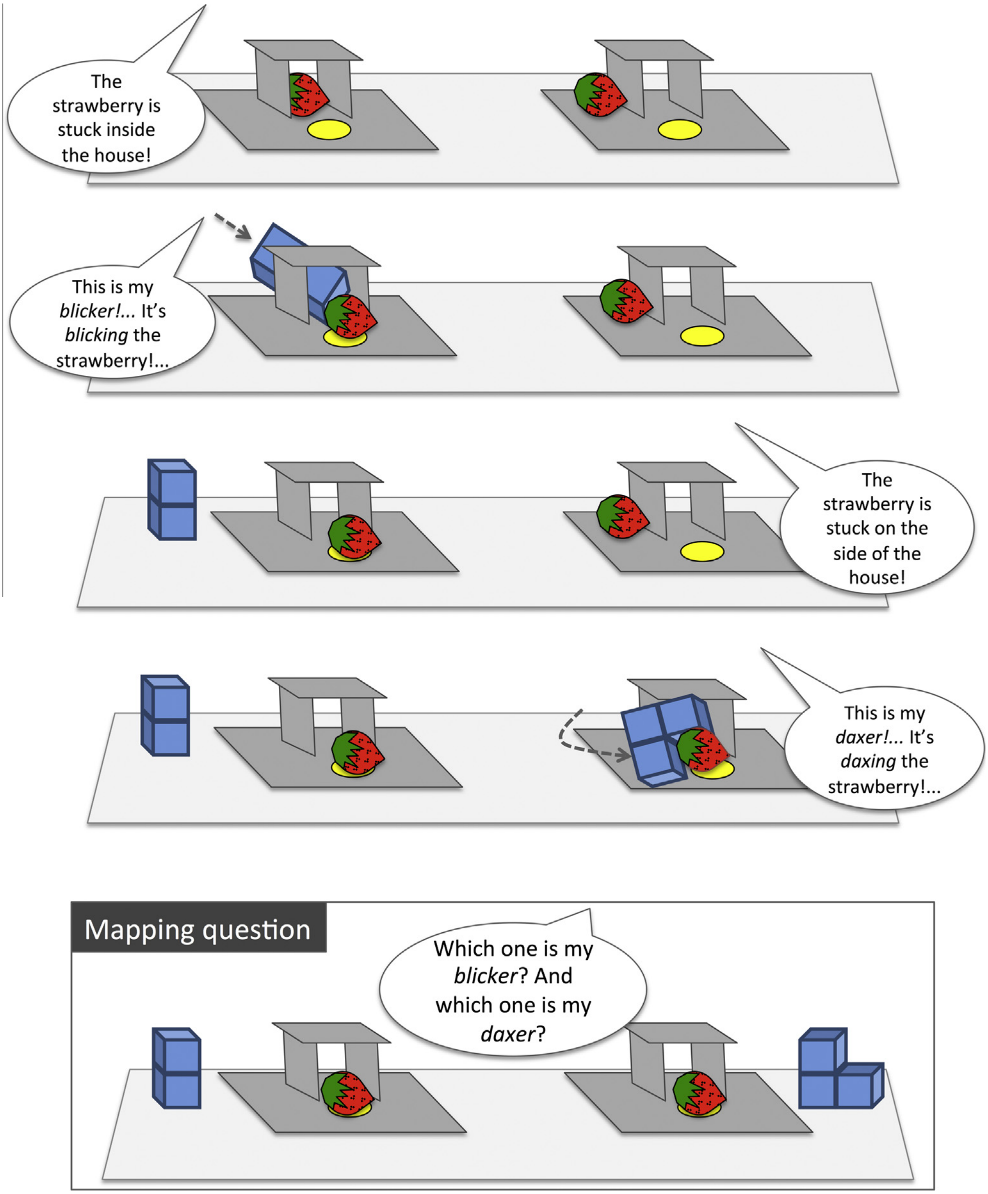

Fig. 1. Teaching Phase 1 in Experiments 1, 2, and 4. Children saw sets of two and three magnetic blocks conjoined to make tools, saw how the tools could be used, and were taught novel labels for the tools.

Next the experimenter demonstrated the *daxer* (whether the *blicker* or the *daxer* was demonstrated first was counterbalanced across children). She pointed to the other tunnel and said, "Now see this strawberry over here? I want to put it on my plate too, but it's stuck on the side of this house! What am I going to do? I know – I'm going to use my *daxer*!" She then brought out the other three identical blocks, which were magnetically connected in an L-shape, and said, "See, this is my *daxer*!" To show children that the blocks were separate objects that could be conjoined, she pulled the blocks apart, waved them, and then conjoined them again into the L-shape. She then said, "It's going to *dax* the strawberry!" She used the tool to rake the strawberry forward, past the tunnel and onto the gold sticker, saying, "See, it's *daxing* the strawberry!" She held up the tool, pulled apart the three blocks, conjoined them again, and said, "This is my daxer!" She then placed the daxer next to the platform, with its three pieces still connected into the L-shape (Fig. 1). For all children in Experiment 1, "blicker" referred to the 2-object tool and "daxer" referred to the 3-object tool.

To probe whether children had learned the novel labels, the experimenter then asked, "Do you remember which one is the *blicker*? And which one is the *daxer*?" (Mapping Question, see Fig. 1). The objects remained in their conjoined configurations during this Mapping Question. If children pointed correctly, the experimenter moved to Teaching Phase 2. If children pointed incorrectly, the experimenter said, "This one is the (*blicker*/*daxer*)! Let's play again," and then repeated Teaching Phase 1 (8/24 children saw Teaching Phase 1 repeated in this way). Twenty-one out of 24 children correctly identified both tools (88%, binomial test $p < 0.001$, two-tailed) after either one or two demonstrations. Regardless of whether children correctly identified the tools after a second demonstration, the experimenter moved on to Teaching Phase 2.

2.1.3.2.2. Teaching Phase 2: Tool use. Children's learning of the tool names was further reinforced and probed in Teaching Phase 2, the purpose of which was to give children more experience with the tools, labels, and functions. The experimenter removed the tunnels and strawberries, leaving both tools on the table. The experimenter then placed one of the tunnels in the middle of the table. In one trial, she placed a strawberry inside it and said, "If my strawberry is stuck *inside* this house, and I want to put it on my plate so I can eat it, which tool should I use?" (Fig. 2). If children pointed correctly, the experimenter said, "That's right! My *blicker*!". If children answered incorrectly, the experimenter said, "The *daxer*? No, that doesn't work!" and demonstrated that the L-shaped tool could not fit into the tunnel to push the strawberry out. Children were then allowed to point again. This procedure was then repeated in a second trial with the strawberry on the side of the house: children were asked which tool they should use to move the strawberry onto the plate. The order in which the tools were probed in Teaching Phase 2 matched the order in Teaching Phase 1. If children answered incorrectly for both tools, Teaching Phase 1 was repeated (5/24 children saw the demonstration repeated; of those five, two had seen Teaching Phase 1 repeated once prior to Teaching Phase 2). Twenty-one out of 24 children responded correctly on at least one question when asked which tool should be used to move the strawberry onto the plate (88%, binomial test $p < 0.001$, two-tailed). Of those, 14 chose correctly on both questions.

The experimenter then removed the tunnel from the table and asked children to point to the tools as in Teaching Phase 1—this time in the opposite order from that in Teaching Phase 1, in order to prevent children from simply responding on the basis of temporal order. Twenty-three out of 24 children (96%, binomial test $p < 0.001$, two-tailed) answered correctly for both labels.

2.1.3.3. Tool-hiding Familiarization. Next we conducted two Tool-hiding Familiarization trials in order to demonstrate that the tools could be taken apart, hidden in the box, and removed in unconnected pieces. The experimenter placed the empty box on the table. She showed children the *blicker*, with its two pieces conjoined, and said, "Now I am going to hide my *blicker* in the box!" She took the tool

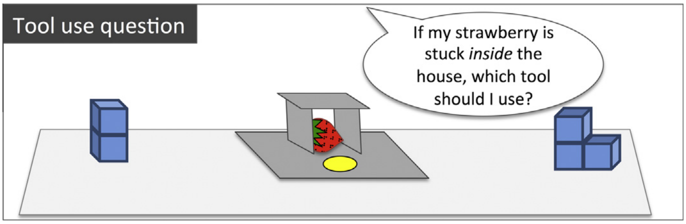

Fig. 2. A sample trial from Teaching Phase 2 in Experiments 1, 2, and 4. Children were asked which tool should be used to move the strawberry onto the plate. The figure depicts a trial in which the 2-block tool is the correct choice. Children were also given a trial in which the 3-block tool was the correct choice.

apart and placed the two blocks on top of the box for children to view, then picked up both blocks and inserted them simultaneously into the box, pushed the box forward, and said, "Can you get my *blicker*?" Children were then allowed to reach in and retrieve both blocks (no blocks were withheld). If children first reached in and retrieved a single block, the experimenter took the block away and placed it out of sight under the table, then allowed children more time to search the box. Once children had retrieved both blocks, the experimenter placed both retrieved blocks out of sight under the table and said, "Great job!" The experimenter then repeated the procedure with the *daxer*. She showed children the *daxer*, with its three pieces conjoined into the L-shape, and said, "Now I am going to hide my *daxer* in the box!" She took the tool apart and placed the three blocks on top of the box, then inserted them simultaneously into the box, pushed the box forward, and said, "Can you get my *daxer*?" Two children failed to retrieve all of the hidden blocks on one of the Tool-hiding Familiarization trials. In those cases, the experimenter said, "Is there anything else in there?" She then reached into the box, retrieved any remaining blocks, and said, "Look what I found!" Blocks were never reassembled into tools or labeled after they were retrieved. The order in which the tools were hidden was counterbalanced across children.

2.1.3.4. Test trials. In the test trials we asked whether children could use the verbal labels alone to infer the number of blocks hidden in the box—that is, whether children could decode the verbal label into the precise number of individual objects in the chunked representation. The experimenter told children that she had hidden either a *blicker* or a *daxer* in the box, out of sight. Children were then allowed to search the box and retrieve two identical blocks. The question was whether children would

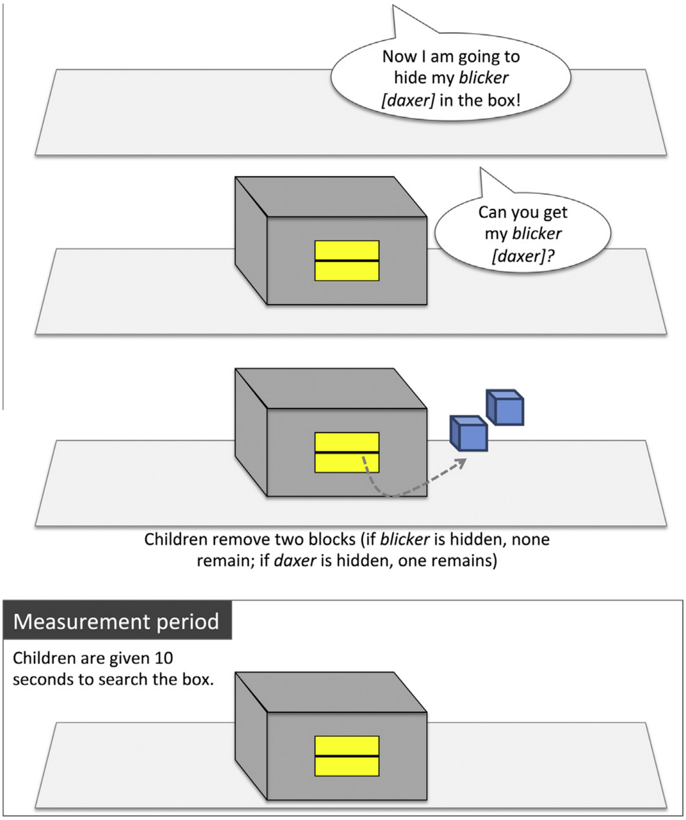

Fig. 3. Test trials from Experiment 1.

continue searching for a third object, and whether this continued searching depended on which verbal label they had heard.

In *None Remaining* trials (Fig. 3), the experimenter showed children the box, shaking it to demonstrate that it was empty. She then moved the box under the table out of children's view and said, "Now I am going to hide my *blicker* in the box!" She brought the box up to eye level, pretended to peek inside, and said, "I see my *blicker*!" She put the box on the table and said, "Can you get my *blicker*?" Children were allowed to reach in and retrieve two unconnected blocks[1]. The experimenter immediately took the two retrieved blocks away and placed them out of view under the table. A 10-s measurement period followed, during which the experimenter looked down to avoid providing any cues, the box was left in place on the table, and children's searching was measured. If children had successfully decoded the label "blicker," they should expect the box to be empty, because they had been taught that a *blicker* was comprised of two objects, had heard that a *blicker* was in the box, and had retrieved two blocks from the box. After 10 s, regardless of children's actions, the experimenter said, "Great job!" and removed the box from the table.

In *More Remaining* trials (Fig. 3), the experimenter showed children the box, shaking it to demonstrate that it was empty. She then moved the box under the table out of children's view and said, "Now I am going to hide my *daxer* in the box!" She brought the box up to eye level, pretended to peek inside, and said, "I see my *daxer*!" She put the box on the table and said, "Can you get my *daxer*?" Children were allowed to reach in and retrieve two unconnected blocks, just as in the *None Remaining* trials. The experimenter immediately took the blocks away and placed them out of view under the table. A 10-s measurement period followed, during which the experimenter looked down, the box was left in place, and children's searching was measured while the experimenter surreptitiously held the third block at the back of the box, out of children's reach, concealed behind the black felt flap so that it would not be visible if children tried to peek. If children had successfully decoded the label "daxer," they should expect the box to contain another object, because they had been taught that a *daxer* was comprised of three objects, had heard that a *daxer* was in the box, but had only retrieved two blocks from the box. After 10 s, regardless of children's actions, the experimenter said, "Let's see what I can find," and reached in and retrieved the missing block. She showed the block to the child and said, "Look what I found!" The experimenter then said "Great job!" and removed the box from the table. On Test trials the blocks always remained unconnected and were never assembled into tools.

Children received four Test trials: two *None Remaining* (*blicker* hidden) and two *More Remaining* (*daxer* hidden) trials. Trial order was counterbalanced: half of the children received trials in the order *None Remaining, More Remaining, More Remaining, None Remaining*, and half in the order *More Remaining, None Remaining, None Remaining, More Remaining.* The amount of time children spent searching the box during the measurement periods (with their hand inserted at least up to the second knuckle) was coded offline by two independent observers using Preferential Looking Coder 1.3 (Libertus, 2011). Inter-observer agreement averaged 95%.

2.1.3.5. What's On This Card?. Besides using recoded representations of object chunks to infer the box's contents, children also could have used explicit numerical representations. That is, children could have associated *blicker* with the count word "two" and *daxer* with "three," but only if they had access to representations of these exact cardinalities. As we were primarily interested in characterizing the abilities of children who could not have verbally counted to solve the task, we measured all children's number word knowledge using the "What's On This Card" (WOC) task (Gelman, 1993; Le Corre & Carey, 2007). We used four sets of cards, each printed with pictures of one to six identical objects (apples, boats, cats, or ladybugs). We showed children the cards in each set, one at a time, always starting with a card depicting a single object, and asked, "What's on this card?" Children often replied with the name of the basic level kind (e.g., "Apple!"), to which the experimenter responded, "That's right, it's ONE apple!" This response has been found to promote counting in children of this age, with

---

[1] Two children had difficulty retrieving the second block one or more trials, because the block had been knocked into one of the box's corners during their first reach. On these trials the experimenter said, "Let's see what I can find," and reached into the front of the box and removed the block. She showed the block to the child and said, "Look what I found!" and then placed the block under the table out of view. The rest of the trial proceeded as described.

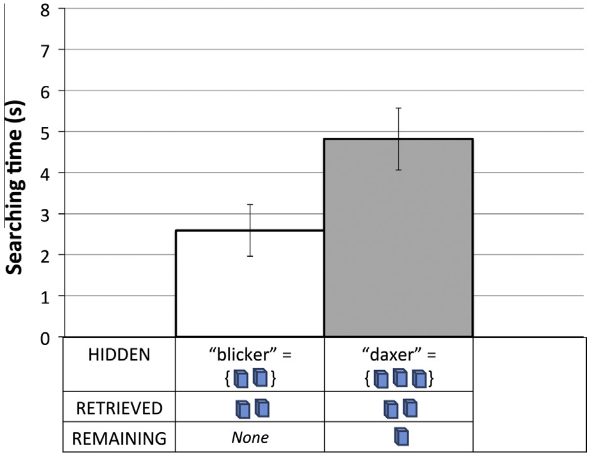

Fig. 4. Mean searching time (in seconds) for *None Remaining* trials and *More Remaining* trials in Experiment 1. Error bars show ±1 SEM. Note that on these trials, children never saw any objects being hidden in the box; rather, they heard that one of the tools was hidden in the box and had to infer how many individual objects were hidden using the tool labels only.

children spontaneously producing number words for the remaining cards (Gelman, 1993). After children had seen the card with one object, they were tested with the other cardinalities in the set (e.g., two, three, four, five, and six apples). The cards within each set were presented in pseudo-randomized numerical order; cards containing one object were always presented first, followed by cards containing two and three objects (in counterbalanced order across sets), followed by cards containing four, five, and six (in counterbalanced order across sets).

Children were asked to produce both a counting response and a cardinality response for each card. For example, if children responded to the prompt of "What's on this card?" with the response "four apples," the experimenter then said, "Can you show me how it's four?" If children responded to the prompt by counting without stating a cardinality, the experimenter said, "So how many apples are there?" The experimenter proceeded through each set of cards until children were unable to produce both a counting and a cardinality response for three cards in a row, at which point the experimenter began the next set of cards. The highest number for which children could produce both responses on at least three out of the four sets of cards was taken to be the child's "knower-level." As in previous work, children who successfully produced both counting and cardinality responses for all of the cards in each set were taken to be Counting Principle-knowers (CP knowers), meaning that they had acquired an understanding of how the verbal count list maps on to sets of objects (e.g. Wynn, 1992). Thus, children in our study were classified as 1-, 2-, 3-, or CP knowers [2]. Previous work suggests that children typically become CP knowers between the ages of 3.5 and 4.5 (e.g., Wynn, 1992); therefore we expected that the younger children we tested here would include many children who did not yet know the precise meanings of number words and who could not count to establish a representation of explicit cardinality.

2.2. Results

2.2.1. What's On This Card (WOC) task

Based on their performance on the WOC task, children were classified as either 1-, 2-, 3-, or CP-knowers. In our sample, 8 children were classified as 1-knowers (6 girls; mean age: 31 months 21 days; range: 24 months 17 days to 34 months 28 days), 8 children as 2-knowers (2 girls; mean

---

[2] Although other researchers have found evidence for 4-knowers (children who know the exact meanings of "one," "two," "three," and "four," but do not yet know the meanings of the other words in their count sequence), we did not observe any 4-knowers in our sample.

age 35 months 4 days; range: 30 months 2 days to 36 months 27 days), 2 children as 3-knowers (2 girls; mean age: 32 months 14 days; range: 30 months 16 days to 34 months 10 days), and 6 children as CP-knowers (3 girls; mean age: 36 months 27 days; range: 36 months 3 days to 38 months 5 days).

2.2.2. Test trials

Our critical question was whether children could decode the novel verbal labels we had taught them, inferring the number of blocks in the box from the label alone. To address this question we examined children's performance with a 2 (Trial Type: *None Remaining* versus *More Remaining*) × 2 (Trial Pair: 1st or 2nd presentation of the trial type) repeated-measures ANOVA. This revealed a significant main effect of Trial Type ($F_{1,23} = 9.38$, $p = 0.006$, $\eta^2 = 0.29$), with children searching longer on *More Remaining* trials (mean = 4.82 s) than on *None Remaining* trials (mean = 2.59 s). That is, after children had retrieved two identical blocks from the box, they continued searching longer if they had previously been told that the box contained a *daxer* than if they had been told it contained a *blicker* (Fig. 4). We found no effect of Trial Pair ($F_{1,23} = 0.11$, $p = 0.74$, $\eta^2 = 0.005$) and no Trial Type × Trial Pair interaction ($F_{1,23} = 2.09$, $p = 0.16$, $\eta^2 = 0.083$); children's performance did not differ from the first to the second presentation of a given trial type.

Next we asked whether children's success was driven by the subset of children who were proficient counters, and who therefore might have verbally counted to represent the number of blocks in each tool. As knowing the meanings of number words "three" and greater might have helped children in our task, we divided children into a group of poorer counters (1- and 2-knowers) and more advanced counters (3- and CP-knowers). We then conducted a 2 (Trial Type: *None Remaining* or *More Remaining*) × 2 (Trial Pair: *1st* or *2nd* presentation of the trial type) repeated-measures ANOVA, with Knower-level (1- and 2-knowers versus 3- and CP-knowers) as a between-subjects factor. This again revealed a significant main effect of Trial Type ($F_{1,22} = 9.61$, $p = 0.005$, $\eta^2 = 0.304$), with children searching longer on *More Remaining* trials than *None Remaining* trials. There was no main effect of Trial Pair ($F_{1,22} = 0.09$, $p = 0.77$, $\eta^2 = 0.004$), and no interaction between Trial Type and Trial Pair. Most importantly, there was no main effect of Knower-level ($F_{1,22} = 0.02$, $p = 0.90$, $\eta^2 = 0.001$), nor any interaction between Knower-level and Trial Type ($F_{1,22} = 0.53$, $p = 0.48$, $\eta^2 = 0.023$). Even children who could not count to determine the cardinality of sets greater than one or two successfully kept searching for a third object after being told that a *daxer* had been hidden (11 out of 16 poorer counters showed this pattern).

As a further check to ensure that partial counting ability did not underlie children's success in Experiment 1, we also examined the performance of children who only showed evidence of understanding the word "one," versus children who knew the meanings of "two" and higher. We again analyzed children's searching time in a 2 (Trial Type: *None Remaining* or *More Remaining*) × 2 (Trial Pair: 1st or 2nd presentation of the trial type) repeated-measures ANOVA, with Knower-level (1-knowers versus 2-, 3-, and CP-knowers) as a between-subjects factor. We again observed a main effect of Trial Type ($F_{1,22} = 10.06$, $p = 0.004$, $\eta^2 = 0.314$) and no interaction between Trial Type and Trial Pair ($F_{1,22} = 2.13$, $p = 0.16$, $\eta^2 = 0.088$). As in the previous analysis, we found no main effect of Knower-level ($F_{1,22} = 0.23$, $p = 0.63$, $\eta^2 = 0.011$) and no interaction between Trial Type and Knower-level ($F_{1,22} = 0.05$, $p = 0.83$, $\eta^2 = 0.002$). Thus, even children who did not know the meanings of any number words other than "one" successfully recoded the sets of blocks using the labels *blicker* and *daxer*, and then decoded these labels to retrieve the precise number of individual objects making up each tool.

2.3. Discussion

Previous studies have found that adults can increase memory efficiency by recoding individual items hierarchically, labeling the resulting chunks, and storing these new representations in long-term memory. To retrieve the individual items, adults then decode the labels, effectively "unchunking" the chunks into their constituents. The results of Experiment 1 suggest that these same basic processes are available by the time children are two to three years old. After learning that an object made up of two identical blocks stacked together was called a *blicker*, and that an object made up of three identical blocks in an L-shape was called a *daxer*, children were able to use the words alone to determine how many objects should be present, with no visual array in sight. That is, after hearing that there was a *daxer* in the box and retrieving only two blocks, children searched the box longer than after being told there was a *blicker* in the box and retrieving two blocks. These results are distinct from previous work showing that even infants can chunk using visual cues (e.g., Feigenson & Halberda, 2004) because previous work showed only that infants can bind representations of individuals into higher order units (and thereby store more information). Children in Experiment 1 formed such chunks, but also bound verbal labels to these new, higher order representations. They then used the labels alone to "unpack" the chunks into their constituents, in the absence of visual input. Further, we showed that this pattern of performance could not be explained by children having simply mapped the words "blicker" and "daxer" onto the already-known words "two" and "three." Even children who did not yet know the meanings of these number words demonstrated that they knew how many blocks were in the box.

However, before accepting the conclusion that the processes of recoding and decoding are available early in development, we must first consider other ways that children might have succeeded at the task. One possibility is that, rather than representing the precise number of blocks comprising each tool, children simply learned that "daxer" meant "more" and "blicker" meant "less." If so, then when children continued to search on trials in which the *daxer* had been hidden, it may not have been because they knew that *daxers* are made up of three objects, whereas only two objects had been retrieved from the box. Instead, children may have kept searching because they had learned that more searching activity typically followed the word "daxer" (for example, because in the Tool Hiding Familiarization that preceded the Test Trials, the experimenter demonstrated more searching following the word "daxer" than the word "blicker").

To rule out this possibility, and to replicate the results of Experiment 1, we tested a new group of children in a version of the task in which correct searching behavior would yield the reverse searching pattern from that of Experiment 1.

3. Experiment 2

In Experiment 2 we sought to replicate the results of Experiment 1, while also testing whether children simply searched more after hearing the word that referred to the three-block tool than after hearing the word that referred to the two-block tool (without representing the individual objects comprising the tools). The Teaching and Tool Hiding Familiarization phases proceeded as in Experiment 1. Children were introduced to each tool and its function, saw that each tool was made of either two or three separable identical blocks, saw that the tools could be taken apart, and saw that the constituent blocks could be hidden in and then retrieved from the box. However, in the Test trials, we reversed the number of objects children were able to retrieve prior to the critical measurement period, relative to that in Experiment 1. On the Test trials in which the experimenter told children that she had hidden the three-block tool, children were now allowed to retrieve *all three* blocks (instead of the two blocks they retrieved in Experiment 1). Hence, for children in Experiment 2, the box should now be expected to be empty. On the Test trials in which the experimenter told children that she had hidden the two-block tool, children were allowed to retrieve *only one* block—hence the box should be expected to contain another object. As such, the searching pattern that would indicate successful memory for the number of blocks in each tool was reversed from that of Experiment 1. Children should now search less during the measurement period when the three-block tool had been hidden than when the two-block tool had been hidden.

An additional goal of Experiment 2 was to explore any effects of the particular novel labels presented. Therefore, half of the children in Experiment 2 were taught that the two-block tool was called a "blicker" and that the three-block tool was called a "daxer" (as in Experiment 1), and the other half of children were taught that the two-block tool was called a "daxer" and that the three-block tool was called a "blicker."

3.1. Method

3.1.1. Participants

Eighteen healthy, full-term children participated [3] (mean age: 33 months 15 days; SD: 2 months 28 days; range: 28 months 9 days to 37 months 25 days; 10 girls). Three children were identified by their parents as black, 13 as white (of whom one was identified as Hispanic or Latino), one as American Indian and black, and one as "other." Of the 97% of parents who reported their educational history, 77% had completed a college degree or higher, 17% had completed some college, 3% had completed high school, and 3% had completed some high school. Six additional children were excluded due to refusal to participate in the task (4), experimenter error (1), or parental interference (1).

3.1.2. Materials and procedure

The materials were identical to those in Experiment 1. For the Generalization trials (see below) we also used arrays of identical green pom-poms and arrays of identical blue poker chips.

The study involved the same three phases as Experiment 1 (a two-part Teaching Phase, a Test Phase, and a Knower-level ("What's on this Card") Phase), along with a new Generalization phase. Teaching Phase 1 was identical to that of Experiment 1, except that half of the children were taught that the two-block tool was called the "blicker" and the three-block tool was called the "daxer," and the other half were taught the reverse. As in Experiment 1, if children responded incorrectly to the Mapping Question at the end of Teaching Phase 1, Teaching Phase 1 was repeated once (7/18 children saw this phase repeated). Seventeen out of 18 children correctly pointed to both requested tools (94%, binomial test $p < 0.001$, two-tailed) after seeing one or two demonstrations. In Teaching Phase 2, when asked to choose which tool was appropriate to retrieve the strawberry, 16 out of 18 children correctly chose for at least one tool (89%, binomial test $p = 0.001$, two-tailed) and 10 of 18 chose correctly for both tools. Five out of 18 children saw Teaching Phase 2 repeated once. When asked a second time to indicate which tool matched a spoken label, 15 out of 18 chose correctly (83%, binomial test $p = 0.008$, two-tailed).

The Test Phase proceeded as in Experiment 1, except that on *None Remaining* trials the experimenter told children that she had hidden the tool made up of three objects (either the *blicker* or the *daxer*, depending on which labeling condition children had been assigned to), and children were allowed to retrieve all three objects from the box. A measurement period then followed, in which we recorded any further searching. On *More Remaining* trials, the experimenter told children that she had hidden the tool made up of two objects, and children were allowed to retrieve only one block (with the second block secretly withheld at the back of the box, using the same method as in Experiment 1); a measurement period then followed. Searching times were later coded offline by two independent observers, and inter-observer agreement averaged 96%.

For 14 out of the 18 children who participated, two additional Generalization trials were added after the Test Phase. These were included in order to ask whether children extended the words "blicker" and "daxer" to other sets of two and three objects, as opposed to restricting the labels to the sets of blue blocks used throughout the experiment. In each Generalization trial, children were shown two sets of new objects: identical green pom-poms on one trial and identical blue poker chips on the other. On each trial, the experimenter placed the objects on the table in two groups spaced approximately 50 cm apart; one group contained two unconnected objects and the other contained three. The experimenter then asked children, "Which one is my 'blicker' / 'daxer?'" Whether the *daxer* was requested first or second, whether the children saw the pom-poms first or second, and whether the two-object group appeared on the left or the right was counterbalanced across children.

Finally, the "What's On This Card" (WOC) task was administered following the Generalization trials.

---

[3] In both Experiments 1 and 2, we were primarily interested in the abilities of children who could not yet count to three and beyond. The testing completed in Experiment 1 allowed us to better predict the ages at which the What's on this Card task would identify such children. Therefore, fewer total children were tested in Experiment 2, yielding numbers of 1- and 2-knowers comparable to those in Experiment 1.

3.2. Results

3.2.1. What's On This Card (WOC) task

As in Experiment 1, children were classified as 1-, 2-, 3-, or CP-knowers based on their performance on the What's On This Card task. There were 10 children who were classified as 1-knowers (5 girls; mean age: 32 months 17 days, range: 28 months 9 days to 36 months 22 days), four 2-knowers (3 girls; mean age: 34 months 20 days; range: 32 months 8 days to 26 months 24 days), one 3-knower (a girl; age: 29 months 15 days), and three CP-knowers (1 girl; mean age: 34 months 13 days; range: 33 months 18 days to 37 months 25 days). As in Experiment 1, children were divided into two groups for the purposes of analysis: poorer counters (1- and 2-knowers; $n = 14$, mean age: 33 months 5 days) and more advanced counters (3- and CP-knowers; $n = 4$, mean age: 34 months 13 days). There were no 4-knowers.

3.2.2. Test trials

First, to ask whether children successfully based their searching on the number of objects comprising each tool, we conducted a 2 (Trial Type: *None Remaining* or *More Remaining*) × 2 (Trial Pair: 1st or 2nd presentation of the trial type) repeated-measures ANOVA. As in Experiment 1, this revealed a significant main effect of Trial Type ($F_{1,17} = 20.22$, $p < 0.001$, $\eta^2 = 0.54$). Children searched longer after being told that the two-block tool was in the box and retrieving just one block (*More Remaining* trials, mean = 5.93 s) than after being told that the three-block tool was in the box and retrieving all three blocks (*None Remaining* trials, mean = 2.75 s) (Fig. 5). There was also a main effect of Trial Pair ($F_{1,17} = 5.38$, $p = 0.03$, $\eta^2 = 0.240$), with children searching longer overall on the first trial pair than the second. There was no interaction between Trial Type and Trial Pair ($F_{1,17} = 0.39$, $p = 0.54$, $\eta^2 = 0.022$).

As in Experiment 1, we next asked whether this success was driven by children who could have counted to determine the number of blocks comprising each tool. To ask whether children's performance differed depending on their counting abilities, we conducted a repeated-measures ANOVA with Trial Type (*None Remaining* or *More Remaining*) and Trial Pair (1st or 2nd) as within-subjects factors and Knower-level (1- & 2-knowers versus 3- and CP-knowers) as a between-subjects factor. This revealed a main effect of Trial Type ($F_{1,16} = 27.41$, $p < 0.001$, $\eta^2 = 0.631$), with children searching longer on *More Remaining* trials than on *None Remaining* trials, but no main effect of Trial Pair ($F_{1,16} = 2.83$, $p = 0.11$, $\eta^2 = 0.150$) and no interaction between Trial Type and Trial Pair ($F_{1,16} = 0.34$, $p = 0.57$, $\eta^2 = 0.021$). Importantly, as in Experiment 1, there was no main effect of Knower-level ($F_{1,16} = 0.001$,

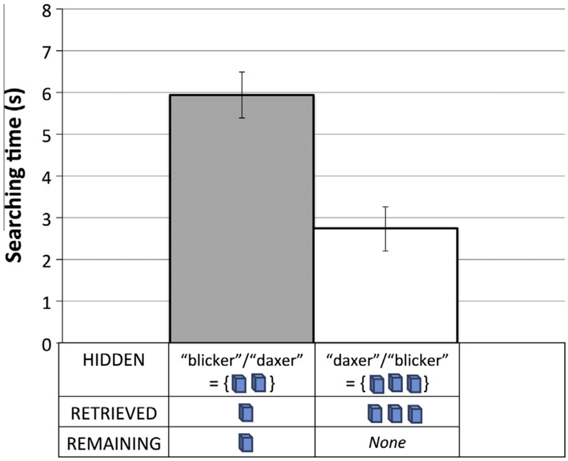

Fig. 5. Mean searching time (in seconds) for *None Remaining* trials and *More Remaining* trials in Experiment 2. Whether the two-block tool was labeled "blicker" or "daxer" was counterbalanced across children. Error bars show ±1 SEM. Note that on these trials, children never saw any objects being hidden in the box, but rather had to infer how many objects were hidden after hearing the labels only.

$p = 0.98$, $\eta^2 < 0.001$), and no interaction between Trial Type and Knower-level ($F_{1,16} = 4.29$, $p = 0.06$, $\eta^2 = 0.212$), suggesting that children who knew the meanings of the words "two" and "three" did not search differently than those who did not know these meanings. Even children who could not count to determine the cardinality of sets greater than one or two successfully represented the number of objects in each tool—they continuing searching after hearing that the two-object tool had been hidden but retrieving only one object, and stopping after hearing that the three-object tool had been hidden and retrieving all three objects (12 out of 14 poorer counters showed this pattern).

As in Experiment 1, we also separated children into 1-knowers versus 2-, 3-, and CP-knowers. Using this separation as a between-subjects factor again yielded a main effect of Trial Type ($F_{1,16} = 21.99$, $p < 0.001$, $\eta^2 = 0.579$), no main effect of Knower-level ($F_{1,16} = 0.003$, $p = 0.96$, $\eta^2 < 0.001$), and no interaction between Trial Type and Knower-level ($F_{1,16} = 1.64$, $p = 0.22$, $\eta^2 = 0.92$). Even children who only knew the meaning of the word "one" showed by their search patterns that they remembered, at least implicitly, the number of blocks comprising each tool.

We also asked whether the tools' names affected children's performance. We conducted a Trial Type × Trial Pair repeated-measures ANOVA with Tool Name ("blicker" = two blocks; "daxer" = three blocks and "daxer" = two blocks; "blicker" = three blocks) as a between-subjects factor. We found no main effect of Tool Name ($F_{1,16} = 0.015$, $p = 0.91$, $\eta^2 = 0.001$) and no interaction between Trial Type and Tool Name ($F_{1,16} = 1.19$, $p = 0.29$, $\eta^2 = 0.069$), suggesting that children's patterns of searching were not influenced by the particular name given to the two-block tool versus the three-block tool.

Finally, we asked whether children generalized the meanings of "blicker" and "daxer" to novel sets of two and three identical objects. Of the children who received generalization trials, 8/14 chose the number-matched group of pom-poms (e.g., if they were taught that "blicker" referred to the two-block tool and "daxer" referred to the three-block tool, they chose the two-object group of pom-poms when asked to point to the *blicker* and the three-object group of pom-poms when asked to point to the *daxer*; 57%, binomial test $p = 0.79$, two-tailed). Nine out of 14 children chose the number-matched group of poker chips (64%, binomial test $p = 0.42$, two-tailed). Hence, we found no evidence that children generalized the words "blicker" and "daxer" to other groups of two or three objects (across pom-poms and poker chips, children answered 17/28 trials correctly (61%), binomial test $p = 0.34$). This suggests that children's representations of *blicker* and *daxer* were specific to the particular objects that had been chunked earlier (i.e., the blue blocks).

3.3. Discussion

Children in both Experiments 1 and 2 demonstrated the ability to recode representations of individual objects into chunks, and to learn novel labels for these chunks. Later, they could decode these labels using knowledge stored in long-term memory, unpacking the chunks into their represented constituents to infer how many objects should be expected in a hiding location. Children's performance was not due to simply mapping the words "blicker" and "daxer" onto the already-known words "two" or "three," since we found no difference between the performance of children who knew the meanings of "two" and "three" and those who did not. Further, Experiment 2 showed that children were not using a strategy of searching more when the larger tool had been hidden, which would not have required them to know exactly how many blocks were in each tool. Rather, our results suggest that children represented the precise number of objects in each chunk. In addition, the results of the Generalization trials of Experiment 2 suggest that children's representations of the individual components of each chunk are restricted to the specific objects that made up the tools, and do not apply to other groups of two and three objects.

Classic studies of recoding in adults demonstrate that recoding increases memory. For example, chess masters can glance at a board, and, as long as the pieces are in legal positions, recode the pieces into groups based on attack and defense positions (Chase & Simon, 1973). These chunked representations enabled experienced players to remember more pieces than novice players could remember. In Experiments 1 and 2, we showed that toddlers also can recode and decode based on learned groupings. However, our results leave open the question of whether this chunking actually can increase the total amount of information children remember. Children in Experiments 1 and 2 were only required to recall two or three objects at a time, since on the critical test trials either a *daxer or* a *blicker* was hidden, but never both. Remembering two or three objects is within the typical working memory capacity observed in adults (e.g., Alvarez & Cavanagh, 2004; Luck & Vogel, 1997) and infants (e.g., Feigenson & Carey, 2003, 2005).

Therefore, we next asked whether the recoding and decoding observed in Experiments 1 and 2 would also enable toddlers to remember more information than they could without chunking. First, in Experiment 3, we confirmed that in the absence of any chunking cues, toddlers would fail to remember five objects. Then, in Experiment 4, we asked whether toddlers could use recoded representations to remember up to five objects in the box, thereby surpassing the typically observed limits on working memory.

4. Experiment 3

In Experiment 3 we sought to confirm that, in the absence of chunking cues, toddlers would fail to represent more than three hidden objects at once. Following the design of previous studies (Barner et al., 2007; Feigenson & Carey, 2003, 2005) we hid different numbers of identical objects in the box (the same blue blocks as in Experiments 1, 2, and 4) and measured children's searching after they had retrieved either all of the hidden objects, or just a subset. We predicted that children would successfully represent two and three hidden objects—i.e., that they would search the box longer after seeing two objects hidden and retrieving just one of them (relative to after seeing two objects hidden and retrieving both), and after seeing three objects hidden and retrieving just two of them (relative to after seeing three objects hidden and retrieving all three). But we predicted that children would fail to represent five hidden objects—i.e., that they would fail to search the box longer after seeing five objects hidden and retrieving just four of them (relative to after retrieving all five).

4.1. Methods

4.1.1. Participants

Nineteen healthy, full-term children participated (mean age: 30 months 28 days; SD: 3 months 13 days; range: 26 months 28 days to 37 months 29 days; 9 girls). Four children were identified by their parents as black, 14 as white, and one as Asian. Of the 87% of parents who reported their educational history, 88% had completed a college degree or higher, 6% had completed some college, and 6% had completed high school). Nine additional children were excluded due to refusal to participate in the task (6), experimenter error (2), or parental interference (1).

4.1.2. Materials and procedure

Children received two Familiarization trials, as in Experiments 1 and 2. Because our aim was to confirm previous findings that children who cannot count fail to represent more than three hidden objects concurrently (Barner et al., 2007; Feigenson & Carey, 2003, 2005), children were not given any evidence to promote chunking the objects into higher order representations. As such, children in Experiment 3 received no Teaching trials and did not see the objects connected in order to create "tools."

On each Test trial, the experimenter placed either two, three, or five blocks on top of the box. As she placed each block, she said "Look, a *blicker*!"[4] Individually labeling each object in an array has been shown to facilitate object individuation in infants (Xu et al., 2005); further, we wanted to make sure that any observed difference in children's performance between Experiment 3 (in which no chunking or recoding cues were provided) and Experiment 4 (in which chunking and recoding cues were provided) was not due to children hearing objects labeled in one experiment but not the other. After placing all of the presented blocks on top of the box, the experimenter pointed to the entire array and said, "Look! In we go!" and inserted all of the blocks simultaneously through the front opening. She then pushed the box forward and said, "What can you find?" Children were allowed to reach in and retrieve either all of the blocks, followed by a 10 s measurement period (*None Remaining* trial), or all of the blocks except

---

[4] Half of children heard the individual blocks labeled "blickers" and the other half heard the blocks labeled "daxers."

one [5], followed by a 10 s measurement period (*More Remaining* trial). As in Experiments 1 and 2, the experimenter always looked down and remained silent during the measurement periods.

Children received eight total Test trials, presented in four pairs of one *None Remaining* trial and one *More Remaining* trial. Children received one trial pair in which two blocks were hidden, one pair in which three blocks were hidden, and two pairs in which five blocks were hidden (this allowed us to obtain more measurements for the array size that was of primary interest in this experiment). The order in which these array sizes were presented was counterbalanced across children. For the five-object array, each child received one Test pair in which the *None Remaining* trial came first and one in which the *More Remaining* trial came first. For the two-object array, half the children received the *None Remaining* trial and half received the More Remaining trial first; the same was true for the three-object array.

4.2. Results

We examined children's searching in a 2 (Trial Type: *None Remaining* or *More Remaining*) × 3 (Number Hidden: two, three, or five objects) repeated-measures ANOVA. This revealed a significant main effect of Trial Type ($F_{1,18} = 6.49$, $p = 0.02$, $\eta^2 = 0.265$), with children searching more overall on *More Remaining* trials than on *None Remaining* trials. There was also a main effect of Number Hidden ($F_{2,36} = 13.04$, $p < 0.001$, $\eta^2 = 0.420$), and a significant interaction between Trial Type and Number Hidden ($F_{2,36} = 11.54$, $p < 0.001$, $\eta^2 = 0.391$). We further examined this main effect and interaction with a series of planned paired-samples t-tests, which revealed that when either two or three objects had been hidden, children searched longer on *More Remaining* than on *None Remaining* trials (two objects hidden: $t_{18} = 6.20$, $p < 0.001$, two-tailed; three objects hidden: $t_{18} = 2.07$, $p = 0.05$, two-tailed). But when five objects had been hidden, children failed to show this pattern—they did not search any longer when another object was expected in the box than when the box was expected to be empty ($t_{18} = 1.70$, $p = 0.11$).

4.3. Discussion

Experiment 3 confirmed that young children are unable to maintain representations of five individual objects concurrently—a finding consistent with previous estimates of working memory capacity for both infants (e.g., Feigenson & Carey, 2003, 2005) and adults (e.g., Alvarez & Cavanagh, 2004; Luck & Vogel, 1997). Next, in Experiment 4, we asked whether recoding would allow children to successfully remember five hidden objects. Children in Experiment 4 received Teaching Trials as in Experiments 1 and 2; these were designed to teach them the names and functions of the two tools that could be created out of different numbers of constituent blocks. This time, in the Test trials children were told that the experimenter had hidden either the *blicker,* the *daxer,* or the *blicker* and *the daxer* in the box. Successfully remembering the number of blocks comprising the *blicker* and *the daxer* would require concurrently representing five identical objects, and therefore would show that children can use verbal recoding to increase the number of hidden objects they can maintain memory.

5. Experiment 4

5.1. Methods

5.1.1. Participants

Eighteen healthy, full-term children participated (mean age: 32 months 15 days; SD: 3 months 1 day; range: 28 months 26 days to 37 months 25 days; 11 girls). Two children were identified by their parents as black (of whom one was identified as Hispanic or Latino), 15 as white, and one as Asian and

---

[5] In order to equate the duration between the hiding of objects and their retrieval across the different array sizes, the experimenter occasionally "assisted" children in retrieving objects. Children generally retrieved the first few blocks themselves, and the experimenter helped children retrieve the others (either all of the others, on *None Remaining* trials, or all but one, on *More Remaining* trials).

white. Of the 92% of parents who reported their educational history, 88% had completed a college degree or higher, 6% had completed some college, and 6% had completed high school). Five additional children were excluded due to refusal to participate in the task (3) or experimenter error (2).

5.1.2. Apparatus and procedure

The Familiarization and Teaching trials proceeded just as in Experiments 1 and 2. As in Experiment 2, whether the two-object tool was labeled "blicker" or "daxer" was counterbalanced across children. As in Experiments 1 and 2, children successfully learned the names of the tools when the tools were in their conjoined configurations. In Teaching Phase 1, 16 out of 18 children correctly pointed to both requested tools (89%, binomial test $p = 0.001$, two-tailed) (4 out of 18 saw the trial repeated once). In Teaching Phase 2, when asked to choose which tool should be used to retrieve the strawberry, 16 out of 18 children correctly chose for at least one tool (89%, binomial test $p = 0.001$, two-tailed) and 10 of 18 chose correctly for both tools (8 out of 18 saw the trial repeated once). When asked again to indicate which tool matched a spoken label, 17 out of 18 children chose correctly (94%, binomial test $p < 0.001$, two-tailed).

During the Test trials, the experimenter hid the tools in the box out of children's view, as in Experiments 1 and 2. She then told children that she was hiding either the *blicker* or the *daxer,* as in Experiments 1 and 2, or that she was hiding the *blicker* and *the daxer.* Children were then allowed to retrieve either all of the blocks (*None Remaining* trials), or all but one block (*More Remaining* trials). As children removed each block from the box, the experimenter took it and placed it on the table next to the box in a random configuration, such that no blocks formed tools [6,7]. Once the requisite number of blocks had been retrieved, the experimenter cleared all of the blocks and placed them out of sight under the table. Children were then allowed 10 s to search the box, during which the experimenter looked down to avoid providing any cues. On *More Remaining* trials, after the 10-s measurement period had ended the experimenter said, "Let's see what I can find!" and then reached into the front of the box and retrieved the missing block. She showed children the block and said, "Look what I found!" The experimenter then said "Great job!" and removed the box from the table. On *None Remaining* trials, after the 10-s measurement period, the experimenter simply said "Great job!" and removed the box from the table.

Children received eight Test trials presented in four pairs. There were two pairs of trials in which only one tool was hidden (One Tool trials). Each One Tool pair contained one *None Remaining* trial, in which just a single tool (either the *blicker* or the *daxer)* was hidden and all of the blocks retrieved (i.e., children were told that the two-block tool had been hidden, and were able to retrieve two blocks, or children were told that the three-tool block had been hidden, and were able to retrieve three blocks), and one *More Remaining* trial, in which the same single tool was hidden and all blocks but one were retrieved (i.e., children were told that the two-block tool had been hidden, and were able to retrieve just one block, or children were told that the three-block tool had been hidden, and were able to retrieved just two blocks). The order in which the *None Remaining* and *More Remaining* trials were presented within each One Tool trial pair was counterbalanced; if children received the *None Remaining* trial first in the first One Tool trial pair, they received the *More Remaining* trial first in the second One Tool trial pair.

There were also two trial pairs in which both the *blicker* and the *daxer* were hidden (Two Tool trials). Each Two Tool trial pair contained one *None Remaining* trial, in which both the *blicker* and the *daxer* were hidden and all of the blocks were retrieved (i.e., children were told that both the two-block tool and the three-block tool had been hidden, and were able to retrieve five blocks), and one *More Remaining* trial, in which both tools were hidden and all blocks but one were retrieved (i.e., children were told that both the two-block tool and the three-block tool had been hidden, and were able to

---

[6] Because the test trials could involve retrieving multiple hidden objects (up to five), we reduced the working memory demands of the task by allowing children to see the blocks they had already retrieved. This is analogous to the chess masters in the classic studies of expert memory, who are permitted to view the pieces they have already placed when attempting to reconstruct the board from memory (e.g., Chase & Simon, 1973).

[7] As in Experiment 3, the experimenter often assisted children in retrieving objects in order to equate the retention interval across set sizes.

retrieve just four blocks). As in the One Tool trial pairs, the order in which the *None Remaining* and *More Remaining* trials were presented in each pair was counterbalanced across trial pairs.

The order in which trial pairs were presented was counterbalanced across children—children saw either One Tool, Two Tools, One Tool, Two Tools or they saw Two Tools, One Tool, Two Tools, One Tool. Whether the tool hidden in the first One Tool trial pair was the *blicker* or the *daxer* was counterbalanced across children. For the Two Tool trial pairs, the experimenter counterbalanced the order in which she named the two tools (e.g., for one trial pair she said, "Now I am going to hide my *blicker* and my *daxer*!" and for the other she said, "Now I am going to hide my *daxer* and my *blicker*!"). Searching times were later coded offline by two independent observers, and inter-observer agreement averaged 98%.

After the Test trials, all children received two Generalization trials, as in Experiment 2, and then were tested on the "What's On This Card" (WOC) task to determine their number word knower-level.

5.2. Results

5.2.1. What's On This Card (WOC) task

As in Experiments 1 and 2, children were classified as 1-, 2-, 3-, or CP-knowers based on their performance on the What's On This Card task. There were seven children classified as 1-knowers (3 girls; mean age: 32 months 28 days; range: 28 months 26 days to 37 months 25 days), five 2-knowers (3 girls; mean age: 32 months 24 days; range: 30 months 6 days to 36 months 15 days), four 3-knowers (3 girls; mean age: 32 months 20 days, range: 29 months 12 days to 35 months 5 days) and two CP-knowers (2 girls; mean age: 34 months 10 days; range: 31 months 20 days to 37 months 0 days). There were no 4-knowers. We again divided children into two groups for further analysis: poorer counters (1- and 2-knowers; $n = 12$, mean age: 32 months 26 days) and more advanced counters (3- and CP-knowers; $n = 6$, mean age: 33 months 6 days).

5.2.2. Test trials

Our primary question of interest was whether children could infer the precise number of objects in the box when told that a *blicker* had been hidden, when told that a *daxer* had been hidden, and when told that a *blicker and a daxer* had been hidden together. We conducted a 2 (Trial Type: *None Remaining* or *More Remaining*) × 3 (Tool(s) Hidden: two-block tool, three-block tool, or both tools) repeated-measures ANOVA. As in Experiments 1 and 2, we found a main effect of Trial Type ($F_{1,17} = 27.55$, $p < 0.001$, $\eta^2 = 0.618$): in general, children searched longer on *More Remaining* than *None Remaining* trials. There was no main effect of Tool(s) Hidden ($F_{2,34} = 2.94$, $p = 0.07$, $\eta^2 = 0.147$), and, crucially, no interaction

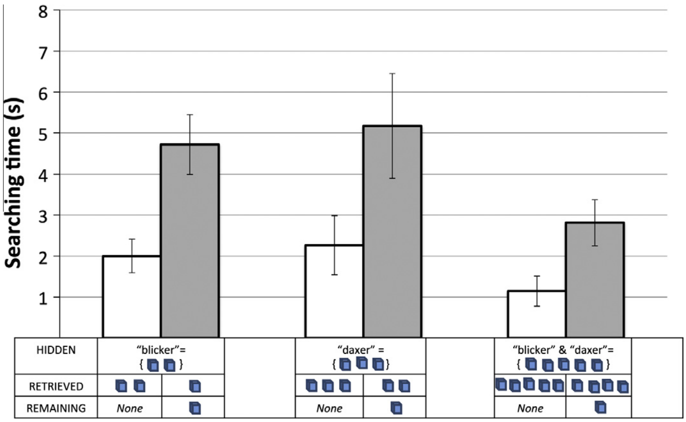

Fig. 6. Mean searching times (in seconds) in Experiment 4. Error bars show ±1 SEM. Note that on these trials, children never saw any objects being hidden in the box, but rather had to infer how many objects were hidden using the labels only.

between Trial Type and Tools(s) Hidden ($F_{2,34} = 0.61$, $p = 0.55$, $\eta^2 = 0.035$). That is, children searched longer on *More Remaining* than *None Remaining* trials when they heard that a *blicker* had been hidden, when they heard that a *daxer* had been hidden, and also when they had heard that both a *blicker* and a *daxer* had been hidden.

Paired *t*-tests confirmed this result. When told that the two-block tool had been hidden, children searched longer after retrieving only one block (mean = 4.72 s) than after retrieving two (mean = 2.0 s) ($t_{17} = 3.34$, $p = 0.004$, two-tailed). When told that the three-block tool had been hidden, children searched longer after retrieving only two blocks (mean = 5.17 s) than after retrieving three (mean = 2.27 s) ($t_{17} = 2.88$, $p = 0.01$, two-tailed). And when told that both tools had been hidden together, children searched longer after retrieving only four blocks (mean = 2.81 s) than after retrieving all five (mean = 1.14 s) ($t_{17} = 2.55$, $p = 0.02$, two-tailed) (Fig. 6).

Next we asked whether this pattern of searching was related to children's counting proficiency. We conducted a 2 (Trial Type: *None Remaining* or *More Remaining*) × 3 (Tool(s) Hidden: two-block tool, three-block tool, or both tools) repeated-measures ANOVA, this time with children's Knower-level (*poorer counters* (1- and 2-knowers) or *more advanced counters* (3- and CP-knowers)) as a between-subjects factor. We again found a significant main effect of Trial Type ($F_{1,16} = 22.83$, $p < 0.001$, $\eta^2 = 0.588$), with children searching longer on *More Remaining* than *None Remaining* trials, but no main effect of Tools(s) Hidden ($F_{2,32} = 2.25$, $p = 0.12$, $\eta^2 = 0.123$) and no interaction between Trial Type and Tool(s) Hidden ($F_{2,32} = 0.35$, $p = 0.70$, $\eta^2 = 0.022$). Importantly, as in Experiments 1 and 2, there was no effect of Knower-level ($F_{1,16} = 0.47$, $p = 0.50$, $\eta^2 = 0.029$) and no interaction between Trial Type and Knower-level ($F_{1,16} = 0.005$, $p = 0.94$, $\eta^2 < 0.001$). Children searched longer on trials when more objects should be expected in the box, even when five objects had been hidden and only four retrieved, regardless of their counting abilities (on two-block tool trials, 11 out of 12 poorer counters showed this pattern; on three-block tool trials, 10 out of 12 poorer counters showed this pattern, and on five-block tool trials, 9 out of 12 poorer counters showed this pattern).

As in Experiments 1 and 2, we also separated children into 1-knowers versus 2-, 3-, and CP-knowers and used this separation as a between-subjects factor. This again yielded a main effect of Trial Type ($F_{1,16} = 25.21$, $p < 0.001$, $\eta^2 = 0.612$) but no main effect of Knower-level ($F_{1,16} = 0.89$, $p = 0.34$, $\eta^2 = 0.053$) and no interaction between Trial Type and Knower-level ($F_{1,16} = 2.95$, $p = 0.11$, $\eta^2 = 0.155$). Even the poorest counters searched more when there was a block remaining in the box than when there were none remaining.

5.2.3. Generalization trials

As in Experiment 2, we asked whether children generalized the meanings of "blicker" and "daxer" to other sets of two and three objects—two versus three identical green pom poms on one trial, and two versus three identical blue poker chips on the other. On both trials the experimenter asked children, "Which one is my *blicker*/*daxer*?" Two children failed to respond on either Generalization Trial. Of the remaining children, 8/16 chose the number-matched set of pom-poms (e.g., if they were taught that "blicker" referred to the two-block tool and "daxer" referred to the three-block tool, they chose the two-object set of pom-poms when asked to point to the "blicker" and the three-object set when asked to point to the "daxer;" 50%, binomial test $p = 1.0$, two-tailed). Only 4/16 children chose the number-matched set of poker chips (25%, binomial test $p = 0.08$, two-tailed). Overall, as in Experiment 2, children did not generalize the words "blicker" and "daxer" to other sets of two or three objects (12/32 trials correct (38%), binomial test $p = 0.22$).

5.3. Discussion

In Experiment 4 we found that toddlers were able to use a verbal label to represent the number of blocks in a hidden array, even when the number of objects in the array exceeded typical working memory limits. This suggests that young children not only can recode groups of objects into chunks and assign these chunks a verbal label, but that these recoded chunks can increase the total amount of information children remember.

6. General discussion

Some 50 years ago, Miller identified a critical series of processes in memory – namely the ability to *recode* to-be-remembered information into higher-order chunks, which could later be precisely *decoded* into their constituents. Miller suggested that through these processes, any limit on the span of immediate memory was easily overcome because links between long-term and working memory could be made rapidly. For example, subject S.F. could accurately recall lists of up to 80 digits by recoding individual digits into semantically meaningful chunks that could later be decoded back into individual digits (Ericsson et al., 1980). This use of recoding and decoding to overcome memory constraints is not limited to expert chess players or to people with extensive laboratory training. As Miller (1956) observed, recoding is largely automatic in adults, and supports everyday remembering. Simple examples like the letter string FBICIAKGB, which adults easily and rapidly recode into three semantically meaningful chunks, illustrate how important this process is for overcoming working memory limits in daily life.

In the present series of experiments we probed the developmental origins of recoding and decoding in memory. In particular, we asked whether toddlers can bind representations of individual objects into higher-order chunks, learn novel labels for the chunks, store these labels in memory, and then later decode the labels to infer the number of objects in a hiding location. In four experiments we demonstrated that toddlers succeeded at this when taught verbal labels for "tools" that could be constructed out of different numbers of identical objects.

In Experiment 1, toddlers successfully used a novel tool label to infer how many objects were hidden in a box. Because the objects children retrieved contained no relevant featural or configural cues to identify them as belonging to one tool versus the other (as they were always identical and unconnected), children's success shows that they represented (at least implicitly) the number of objects comprising each tool, and that they were able to retrieve this information using only the verbal labels. Experiment 2 replicated this result, ruled out an alternative explanation for children's success, and showed that the particular verbal labels given to the two-object and three-object tools did not influence children's success. Finally, in Experiments 3 and 4, we asked whether recoding objects in this way increases the amount of information toddlers can remember, as it does for adults. First, in Experiment 3 we replicated the finding that although children successfully represent arrays of two and three hidden objects, they fail to accurately represent larger arrays (in this case, arrays of five objects). Then, in Experiment 4, we showed that toddlers used their new knowledge of *blickers* and *daxers*—and critically, the number of individual objects needed to construct each of these tools – to represent more objects than they otherwise could. In contrast to those in Experiment 3, children in Experiment 4 successfully remembered not only arrays of two and three hidden objects, but also arrays of five – thereby overcoming the typically observed limits on working memory for unchunked objects (Cowan, 2001; Feigenson & Carey, 2003, 2005; Luck & Vogel, 1997). Furthermore, across all of our experiments we showed that children could not have succeeded simply by mapping the words "blicker" and "daxer" onto the already-known words "two" and "three." Children who did not know the meanings of the words "two" and "three" performed as well as children who knew these number word meanings.

Our findings are distinct from previous work showing chunking by infants and children in several ways. First, previous studies showed that infants can chunk items in working memory using cues such as spatial proximity or shared features (Feigenson & Halberda, 2004; Moher et al., 2012; Rosenberg & Feigenson, 2013), but for this kind of chunking, links to long-term memory are not needed. The chunked object representations can be maintained in working memory over brief durations, supporting enhanced memory performance immediately after the objects are chunked, without being held in any durable long-term store. In contrast, children in the present experiments maintained the chunked representations that they had learned during the training phase over the rest of the testing session; they represented the number of objects in each chunk even after many minutes and multiple intervening trials.

Second, although infants can chunk using knowledge about category membership (Feigenson & Halberda, 2008), social relationships (Stahl & Feigenson, 2014), or temporal regularities between object features (Kibbe & Feigenson, submitted for publication), no previous work had demonstrated that children can represent a *precise* relationship between the knowledge used to chunk and the chunk itself. Indeed, infants' representations of chunks appear to be markedly imprecise; infants often fail to store featural information about the objects that are the chunk's constituents (Kibbe & Feigenson, submitted for publication; Rosenberg & Feigenson, submitted for publication). Similarly, studies of older children (e.g. Cowan et al., 2010; Kobasigawa & Orr, 1973; Rosner, 1971) show that they can increase the amount of information stored in working memory by hierarchically restructuring the information using long-term knowledge about category membership. However, this kind of chunking again lacks a precise recoding and decoding scheme: for example, the higher order category *animal* cannot be precisely decoded into *cat, dog*, and *fish*, because *animal* picks out a large and unspecified number of basic level kinds. In contrast to these cases, a recoded chunk (in Miller's sense) is a memory structure that can *only* be decoded into one particular set of individuals. Our results suggest that toddlers can represent a precise relationship between the recoded chunk and the chunk's constituents. This specificity is further evidenced by our finding that children did not generalize the labels to novel groups of two and three objects—children appeared to decode, e.g., "blicker" into something like "the tool that is made using a block and a block," rather than as an abstract cardinality that could apply to any set of two objects.

Finally, no study has previously shown that young children can assign a verbal label to a chunk of items, and then retrieve representations of the individual constituent items using the verbal label alone. Toddlers' ability to do this allows them increased flexibility. The representation of a chunk need not be actively maintained in working memory, but instead can be offloaded to long-term memory if not presently needed, and later retrieved using a verbal code.

Importantly, we found that this ability to link a verbal label to a higher order, chunked representation was exhibited even by children who could not yet count and who did not know the meaning of number words. Even children who were poor counters quickly learned the meanings of the words "blicker" and "daxer," where each word referred to a tool made from a precise number of component objects. Unlike number words, which refer to the cardinality of any set of objects (e.g., two blocks, three triumphs, four ideas), the words "blicker" and "daxer" could not be applied in this way and children recognized this, as shown by their failure to generalize these labels to sets of two and three novel objects. We suggest that it might therefore be useful to think of "blicker" and "daxer" as "proto-number words." "Proto-number word" representations lack at least two key features of number words. First, they lack the generality of number words. Whereas "three" can refer to the cardinality of a set of objects, sounds, actions, ideas, or mental entities, "daxer" refers only to the tool that can be made from three blocks, and that can be used for a particular function. Second, there is no evidence that children's representations of *blicker* and *daxer* captured information about the ordinal relations between the chunks. Whereas a critical aspect of the meaning of "three" is that it is exactly one more than "two" and exactly one less than "four," *blicker* and *daxer* need not be thought of as ordinally related in any way. However, we suggest that recoded verbal labels (or "proto-number words") share at least one property of actual number words—they reference precise, quantity-relevant information—and this opens avenues for future work, such as investigating the role that chunked and recoded representations might play in the protracted process by which children come to master the meanings of the count words.

In summary, we showed that toddlers can mentally reorganize representations of individual objects into higher order units, learn new words for these units, and later use the words to make inferences about the number of objects in a location. Further, toddlers successfully used this hierarchical structure to remember more objects than is typical given the strict limits of working memory. Our results therefore suggest that the processes of recoding and decoding in memory are available and are spontaneously deployed by the second year of life, suggesting that the computations underlying this important component of memory are continuous across human development.

Acknowledgment

This work was supported by a James S. McDonnell Foundation Scholar Award to L.F.

References

Alvarez, G. A., & Cavanagh, P. (2004). The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychological Science, 15*, 106–111.

Barner, D., Thalwitz, D., Wood, J., Yang, S. J., & Carey, S. (2007). On the relation between the acquisition of singular-plural morph-syntax and the conceptual distinction between one and more than one. *Developmental Science, 10*(3), 365–373.

Booth, A. E., & Waxman, S. (2002). Object names and object functions serve as cues to categories for infants. *Developmental Psychology, 38*(6), 948.

Bower, G. H. (1970). Organizational factors in memory. *Cognitive Psychology, 1*(1), 18–46.

Bower, G. H., Lesgold, A. M., & Tieman, D. (1969). Grouping operations in free recall. *Journal of Verbal Learning and Verbal Behavior, 8*(4), 481–493.

Bower, G. H., & Winzenz, D. (1969). Group structure, coding, and memory for digit series. *Journal of Experimental Psychology, 80*(2p2), 1–38.

Charness, N. (1976). Memory for chess positions: Resistance to interference. *Journal of Experimental Psychology: Human Learning and Memory, 2*(6), 641.

Chase, W. G., & Ericsson, K. A. (1980). Skilled memory. In *Cognitive skills and their acquisition*, Anderson, J. R., (Ed.), Psychology Press.

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology, 4*, 55–81.

Chi, M. T. H., Glaser, R., & Rees, E. (1982). Expertise in problem solving. In R. J. Sternberg (Ed.). *Advances in the psychology of human intelligence* (Vol. 1). Hillsdale, NJ: Erlbaum.

Christie, S., & Gentner, D. (2014). Language helps children succeed on a classic analogy task. *Cognitive Science, 38*(2), 383–397.

Cowan, N. (2001). The magical mystery four: How is working memory capacity-limited, and why? *Current Directions in Psychological Science, 19*, 51–57.

Cowan, N., Hismjatullina, A., AuBuchon, A. M., Saults, J. S., Horton, N., Leadbitter, K., et al (2010). With development, list recall includes more chunks, not just larger ones. *Developmental Psychology, 46*(5), 1119–1131.

De Groot, A. D. (1965). *Thought and choice in chess*. The Hague: Mouton.

Dirlam, D. K. (1972). Most efficient chunk sizes. *Cognitive Psychology, 3*, 355–359.

Epelboim, J., & Suppes, P. (2001). A model of eye movements and visual working memory during problem solving in geometry. *Vision Research, 41*, 1561–1574.

Ericsson, K. A., Chase, W. G., & Faloon, S. (1980). Acquisition of a memory skill. *Science, 208*, 1181–1182.

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review, 102*(2), 211–245.

Feigenson, L., & Carey, S. (2003). Tracking individuals via object-files: Evidence from infants' manual search. *Developmental Science, 6*, 568–584.

Feigenson, L., & Carey, S. (2005). On the limits of infants' quantification of small object arrays. *Cognition, 97*, 295–313.

Feigenson, L., & Halberda, J. (2004). Infants chunk object arrays into sets of individuals. *Cognition, 91*, 173–190.

Feigenson, L., & Halberda, J. (2008). Conceptual knowledge increases infants' memory capacity. *Proceedings of the National Academy of Sciences, 105*, 9926–9930.

Frey, P. W., & Adesman, P. (1976). Recall memory for visually presented chess positions. *Memory & Cognition, 4*(5), 541–547.

Gelman, R. (1993). A rational-constructivist account of early learning about numbers and objects. In D. L. Medin (Ed.). The psychology of learning and motivation. *Advances in research theory* (Vol. 30, pp. 61–96).

Gilchrist, A. L., Cowan, N., & Naveh-Benjamin, M. (2009). Investigating the childhood development of working memory using sentences: New evidence for the growth of chunk capacity. *Journal of Experimental Child Psychology, 104*, 252–265.

Hitch, G. J., Burgess, N., Towse, J. N., & Culpin, V. (1996). Temporal grouping effects in immediate recall: A working memory analysis. *The Quarterly Journal of Experimental Psychology, 49A*(1), 116–139.

Kibbe, M. M. & Feigenson, L. (submitted for publication). Infants use temporal regularities to chunk objects in working memory.

Kobasigawa, A., & Orr, R. R. (1973). Free recall and retrieval speed of categorized items by kindergarten children. *Journal of Experimental Child Psychology, 15*(2), 187–192.

Le Corre, M., & Carey, S. (2007). One, two, three, four, nothing more: An investigation of the conceptual sources of the verbal counting principles. *Cognition, 105*(2), 395–438.

Libertus, K. (2011). Preferential Looking Coder (Version 1.3.3) [Computer software]. Baltimore, MD: Johns Hopkins University.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature, 390*(6657), 279–281.

Mandler, G., Pearlstone, Z., & Koopmans, H. S. (1969). Effects of organization and semantic similarity on recall and recognition. *Journal of Verbal Learning and Verbal Behavior, 8*, 410–423.

Mathy, F., & Feldman, J. (2012). What's magic about magic numbers? Chunking and data compression in short-term memory. *Cognition, 122*, 346–362.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*(2), 81.

Miller, G. (1958). Free recall of redundant strings of letters. *Journal of Experimental Psychology, 56*(6), 485–491.

Moher, M., Tuerk, A. S., & Feigenson, L. (2012). Seven-month-old infants chunk items in working memory. *Journal of Experimental Child Psychology, 112*, 361–377.

Rosenberg, R. D., & Feigenson, L. (submitted for publication). Chunking and the loss of featural resolution in infants' object representations.

Rosenberg, R. D., & Feigenson, L. (2013). Infants hierarchically organize memory representations. *Developmental Science, 16*(4), 610–621.

Rosner, S. R. (1971). The effects of rehearsal and chunking instructions on children's multitrial free recall. *Journal of Experimental Child Psychology, 11*(1), 93–105.

Ryan, J. (1969). Grouping and short-term memory: Different means and patterns of grouping. *Quarterly Journal of Experimental Psychology, 21*, 137–147.

Sargent, J., Dopkins, S., Philbeck, J., & Chichka, D. (2010). Chunking in spatial memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*(3), 576–589.

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell Systems Technical Journal, 27*, 379–423.

Simon, H. A. (1974). How big is a chunk? *Science, 183*(4124), 482–488.

Sloutsky, V. M., & Robinson, C. W. (2008). The role of words and sounds in infants' visual processing: From overshadowing to attentional tuning. *Cognitive Science, 32*(2), 342–365.

Sodian, B., Schneider, W., & Perlmutter, M. (1986). Recall, clustering, and metamemory in young children. *Journal of Experimental Child Psychology, 41*(3), 395–410.

Stahl, A., & Feigenson, L. (2014). Social knowledge facilitates chunking in infancy. *Child Development*. http://dx.doi.org/10.1111/cdev.12217 (online first.

Waxman, S. R., & Braun, I. (2005). Consistent (but not variable) names as invitations to form object categories: New evidence from 12-month-old infants. *Cognition, 95*(3), B59–B68.

Waxman, S. R., & Markow, D. B. (1995). Words as invitations to form categories: Evidence from 12- to 13-month old infants. *Cognitive Psychology, 29*(3), 257–302.

Wilcox, T., & Chapa, C. (2004). Priming infants to attend to color and pattern information in an individuation task. *Cognition, 90*(3), 265–302.

Wortman, P. M., & Greenberg, L. D. (1971). Coding, recoding, and decoding of hierarchical information in long-term memory. *Journal of Verbal Learning and Verbal Behavior, 10*(3), 234–243.

Wynn, K. (1992). Children's acquisition of the number words and the counting system. *Cognitive Psychology, 24*, 220–251.

Xu, F., Cote, M., & Baker, A. (2005). Labeling guides object individuation in 12-month-old infants. *Psychological Science, 16*, 372–377.