ELSEVIER

Applied Developmental Psychology 24 (2003) 739–762

Applied Developmental Psychology

Fast mapping of multiple words: Insights into when "the information provided" does and does not equal "the information perceived"

Krista M. Wilkinson[a,b,*], Erin Ross[b], Adele Diamond[b,*]

[a]*Department of Communication Sciences and Disorders, Emerson College, Boston, MA, USA*
[b]*Center for Developmental Cognitive Neuroscience, University of Massachusetts Medical School, Eunice Kennedy Shriver Center Campus, 200 Trapelo Road, Waltham, MA 02452-6332, USA*

Abstract

A candidate process for explaining the rapid vocabulary acquisition during the preschool years is "fast mapping," children's ability to sketch partial maps of a word's meaning after brief exposure. The present study examines this process for learning multiple words, testing the hypothesis that children's attention to the information critical for quickly mapping multiple words onto their referents depends on the alternatives available when the words are introduced. Fifty-eight 40-month-old children participated in one of two conditions. In both conditions, each trial for novel Word #1 presented a novel object and three familiar ones. The conditions differed in the object choices presented when novel Word #2 was introduced. Although the same information was available to children in both conditions, younger children showed significantly better learning of the new words in the successive condition than in the concurrent condition. Implications of this for age-related differences and for teaching strategies are discussed.

Keywords: Fast mapping; Language acquisition; Lexical; Attention; Problem solving; Deduction; Reasoning; Preschoolers

* Corresponding authors. Department of Communication Sciences and Disorders, Emerson College, 120 Boylston Street, Boston, MA 02116, USA. Tel.: +1-617-824-8288; fax: +1-617-824-8735 (K.M. Wilkinson). Center for Developmental Cognitive Neuroscience, Eunice Kennedy Shriver Center Campus, University of Massachusetts Medical School, 200 Trapelo Road, Waltham, MA 02452, USA. Tel.: +1-781-642-0156; fax: +1-781-642-0202 (A. Diamond).

E-mail addresses: Krista_Wilkinson@emerson.edu (K.M. Wilkinson), adele.diamond@umassmed.edu (A. Diamond).

0193-3973/$ – see front matter
doi:10.1016/j.appdev.2003.09.006

1. Introduction

One of the best-documented phenomena of early language development is the remarkable rate of vocabulary expansion in the preschool years (Carey, 1982). Infants at the outset of lexical development acquire new words laboriously, learning an average of three new words a week; older toddler and preschool-aged children seem to learn many new words daily (e.g., Bates, Dale, & Thal, 1995). One candidate process for explaining this rapid vocabulary acquisition has been termed "fast mapping" (Carey & Bartlett, 1978). Fast mapping refers to children's ability to sketch partial maps of a word's meaning after only brief and oftentimes incidental exposure (Rice, 1989).

Children appear able to use semantic, syntactic, and/or social cues to assist them in the process of narrowing down possible meanings for novel words they hear. One such cue is the contrast between novel words that have no assigned meaning and familiar words for known referents (e.g., Golinkoff, Mervis, & Hirsh-Pasek, 1994; Markman, 1989). In one commonly used task, a participant might be presented with color photographs of a dog, tree, and apple and asked to select the dog ("Can you show me [the] dog?"). Next, color photographs of two well-known items (e.g., dog, tree) might be presented along with a photograph of an unusual or novel object (e.g., a garlic press). The child is asked, "Can you show me [the] dax?" The key behavior of interest is whether or not the child selects the novel item. If so, it suggests the occurrence of fast mapping (Dollaghan, 1985).

1.1. The role of fast mapping in lexical development

Many authors have argued that fast mapping may foster the rapid vocabulary acceleration observed between children's first and second birthdays (e.g., Golinkoff, Hirsh-Pasek, Bailey, & Wenger, 1992; Markman, 1989). Both cross-sectional and preliminary longitudinal studies support this claim. The onset of success in fast mapping paradigms appears at the same approximate time as the onset of the rapid increase in vocabulary size, often at about 18 months of age (Dickinson, 1984; Dollaghan, 1985; Golinkoff et al., 1992; Kagan, 1981; Markman, 1989). In one of the only studies following children over time, Mervis and Bertrand (1994) reported that the onset of rapid vocabulary learning co-occurred regularly with the observation of what they defined as fast mapping (selection of a novel toy in the presence of a novel word). Vocabulary size, rather than chronological age, was most highly associated with fast mapping in their research. The correlation of fast mapping selections and vocabulary size was further substantiated in a study of children with Down syndrome as well as children with hearing loss (Lederberg, Preszbindowski, & Spencer, 2000; Mervis & Bertrand, 1995), despite the fact that those children were chronologically older than the typically developing children with whom they were compared.

The studies cited above have been extremely useful in evaluating the point of emergence of fast mapping. However, the onset of the phenomenon is but one component of a larger picture of how fast mapping might contribute to lexical development. If fast mapping fosters rapid vocabulary learning, as claimed, we must ask not just when fast mapping emerges, but how well children who demonstrate fast mapping use the strategy to learn more than one novel word. It is this question that we take up in the present study.

1.2. How do children know which information they should attend to and remember?

A passenger in a vehicle often attends to much less of the information relevant to finding the way than the vehicle's driver. Why? The information best attended and remembered is the information needed to choose among alternative possible responses, and passengers are often not required to make responses relevant to navigation. Perception is focused by the need to choose and by the set of possible alternatives (Gibson & Pick, 2000; Gibson, 1979; Olson, 1964). The need to choose, whether it is deciding which way to turn or which stimulus to select, alerts an organism to the need to attend to information that will help in making that choice. The set of possible alternatives alerts the organism to what information is relevant.

We began this study with the hypothesis that young children's efficiency in quickly mapping multiple new words onto their referents likely depends on how the alternatives are presented when the new words are introduced. If a child must select a novel item from a set of familiar ones, the child need only encode the most minimal information about the novel item (i.e., that it is unfamiliar). If a child encodes multiple new items this way, the child might well have difficulty later determining which of the new words goes with which of the new items. However, if the first new item learned is included within the set of alternatives when a second new word is introduced, that should alert the child that simply attending to which item is unfamiliar would be insufficient, and more information about the first and second new items now becomes relevant.

1.3. Evaluating how fast mapping contributes to acquisition of more than one word

Most studies of fast mapping have explored how children respond when presented with a single novel word. Although a second novel item has sometimes been presented in an array with a target word as a "distractor" during retention tests (Golinkoff et al., 1992; Mervis & Bertrand, 1994), usually the second novel item does not receive a label of its own. This paradigm is well suited for examining when and how fast mapping emerges in toddlerhood. It may not, however, be the most effective means of evaluating how fast mapping, once in place, allows for rapid acquisition of up to 9 new words a day during the preschool years (cf. Carey, 1982). Studies of multiple word learning are more likely to reveal important insights into the process by which fast mapping leads to rapid vocabulary growth.

Two laboratories have focused on rapid acquisition of multiple new words. Rice and her colleagues introduced what they called the "quick incidental learning" paradigm (Rice, 1990; Rice, Buhr, & Nemeth, 1990; Rice, Oetting, Marquis, Bode, & Pae, 1994). In these studies, a television-viewing format was used to introduce new words within a voice-over narration describing the contents of a short movie. The new words replaced common but general labels; thus, for instance, "trudge" replaced a word like "walk" or "go." Children were exposed to as many as five words in each of several categories (objects, action, attributes). Testing for learning occurred after exposure, with children being asked to select among four pictures when presented with a target word. Children with language impairment retained significantly fewer words than typically developing children matched either chronologically or by mean length utterance (Rice et al., 1990). A second study (Rice et al., 1994) indicated that a greater number of exposures to each new word was necessary for children diagnosed with a language impairment to perform comparably to matched peers.

Wilkinson and her colleagues introduced another set of methods by which two or more new words were taught (Wilkinson & Albert, 2001; Wilkinson & Green, 1998; Wilkinson & Shah, 2003). Wilkinson and Green (1998) explored how individuals with severe developmental disabilities acquired two new words through procedures adapted from the new/known contrast underlying fast mapping. In that study, 10 individuals with moderate to severe mental retardation were exposed to two new words using the novel/known contrast. The total number of teaching exposures was set to 24, for two related reasons. First, Rice et al.'s (1994) work had demonstrated that children with language-specific disorders required greater numbers of exposures to new labels than age- or language-matched peers. Second, the targeted participants in Wilkinson and Green's study had intellectual disabilities concomitant with language limitations. Thus, it was expected that such individuals might need even greater support (in the form of multiple teaching exposures) than the children in Rice and colleagues' sample.

After the exposure/instruction sessions, tests for learning were conducted in a separate session occurring 1–3 days after the final teaching session. In these tests, both novel referents were presented to the participant simultaneously and the participant was asked to select between them on the basis of the spoken word. Multiple learning test trials were presented, and participants' overall accuracy on these trials constituted the data subjected to analysis.

Wilkinson and Green (1998) and Wilkinson and Shah (2003) were interested not just in how well their participants learned two new words, but also whether alterations to the teaching procedures changed learning outcomes. Thus, they contrasted two procedures, the successive and the concurrent introduction procedures. Fig. 1 illustrates examples of exposure trials from each procedure, as well as the testing trials for learning outcomes. The concurrent introduction procedure made use of simple contrasts between target novel objects and highly familiar "baseline" objects. For each new word the contrast items in the array were photographs of well-known items like cat, banana, or tree. After exposures to each novel word–object pair, testing for learning then presented the two novel objects together, to see whether or not the participant had learned the labels.

In the successive introduction procedure, exposure to the first new word/object pair was identical to that in concurrent introduction procedure; the new object was contrasted with well-known objects. After several such trials, the second new object and label were introduced. Importantly, the contrast items present in the array when the *second* novel label was spoken included not just a familiar baseline item, but also the just-labeled item (in this case, the item matching to "rutch"). To select correctly, participants had to mark and attend to the differences in the two novel stimuli. Testing for learning was the same as in concurrent introduction condition.

All participants underwent both procedures (successive and concurrent introduction), learning different/unique pairs of novel word–referent pairs in each one. Significant differences in learning outcomes resulted. Wilkinson and Green (1998) reported that while only 3

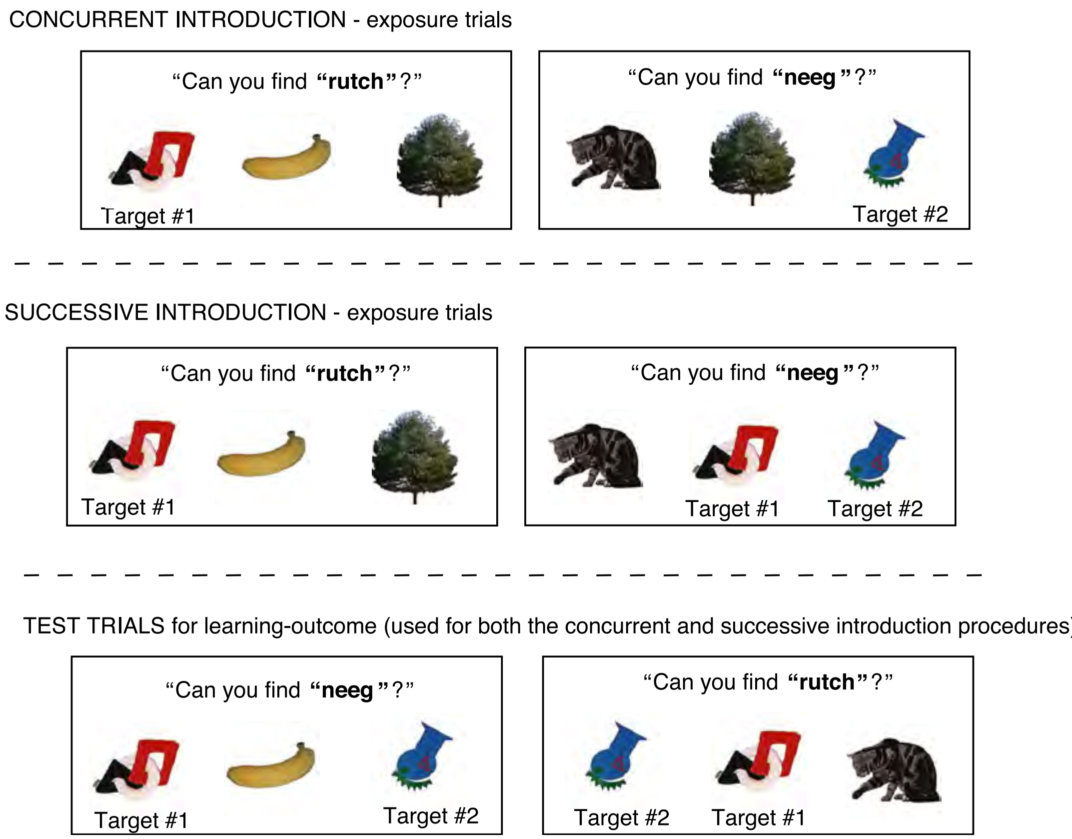

Fig. 1. Examples of trials in concurrent and successive introduction procedures.

of 10 participants learned to distinguish two newly taught words at greater than 85% accuracy under concurrent introduction conditions, 8 of the 10 participants were able to do so under the successive introduction condition. Wilkinson and Shah (2003) examined acquisition of novel words consistent with English phonology (learned within an English word-matching baseline) in typically developing children between the ages of 3 and 5 years who were monolingual or bilingual in English or who were learning English as a second language. They reported that the procedural variations made no difference for children with high familiarity in English (both monolingual and bilingual children performed equally well) but statistically significant differences in outcome emerged in children for whom English was a relatively unfamiliar phonological system. Again, the differences were in favor of the successive introduction procedure.

1.4. Research goals

The present study had three overall goals. One was to refine the methods introduced thus far for examining the relation of fast mapping to rapid vocabulary acceleration. The procedures described in the Wilkinson and Green (1998) and Wilkinson and Shah (2003) studies can only be considered adapted variants from fast mapping, because the number of trials presented (24, distributed across two new words) is higher than one might consider to be "fast" mapping. We were interested in reducing the number of exposures to bring it more into line with the notion of a "fast" map. In the current study, the total number of exposures was reduced from 24 to 6 (3 per word).

A second goal of the study was to examine more carefully the facilitative effect of the successive introduction procedure. The results of prior studies suggested strongly that the successive introduction procedure facilitated learning outcomes even within individual children. Wilkinson and Green (1998) argued that the facilitative effect likely resulted because the successive introduction procedure explicitly required participants to mark the contrast between the second novel item and the first during training, unlike concurrent introduction. However, the methodology they used left open another possible explanation. In concurrent introduction, each novel word–referent pair was introduced in 12 exposure trials before tests for learning. Because successive introduction required a short "practice" with the first novel word–referent pair (so that this referent could then serve as a somewhat familiar contrast for the second new referent), there was an unequal distribution of exposure trials for the first and the second word, with 18 exposure trials for the first new word and only 6 exposure trials for the second novel word. Although the total number of exposures was therefore equal across protocols (total = 24), it is conceivable that the differences in outcomes might be due to differences in exposure distribution. The current study eliminated this confound by providing an equal number of exposures to each new word (3 per word), across both conditions. If the facilitative effect of the successive introduction procedure continues to be demonstrated, it would support the argument that the key alteration in the successive introduction procedure concerned the contrast of the two novel targets (second referent with first referent). Our prediction was that the successive introduction procedure would continue to provide a measure of support for learning, even when the trial numbers were equally balanced for the two words.

The final goal of the study was to examine whether or not differences emerged on the basis of age. If fast mapping is a developmental phenomenon (see Hollich et al., 2000), then we might expect that older children, who should have increased expertise at using the novel/known contrast for accomplishing word learning, would learn easily under either introduction condition. Younger children, who are still in the process of learning how to apply fast mapping for vocabulary learning, would be more likely to benefit from the support offered by the successive introduction procedure. We predicted that procedural variations would have an effect for younger but not older children.

2. Method

2.1. Participants

A total of 58 children participated. Three other protocols were terminated early due to restlessness (2 participants) or experimenter error (1 participant). Overall, 29 of the 58 (50%)

children were female. All children were normally developing according to parent report, and extensive informal cognitive testing for related research did not reveal any cause for concern about their development.

Participants were assigned to successive or concurrent introduction conditions. Twenty-one children underwent concurrent introduction, and 37 underwent successive introduction. The mean age of both groups was 40.2 months, with a range of 26–52 months for the concurrent condition and 29–57 months for the successive condition. Because we had predicted that the successive introduction procedure would facilitate learning, especially in younger children, children were divided into two age groups: those under 42 months (the younger cohort) and those over 42 months (the older cohort). Participant characteristics of each of these cohorts are summarized in Table 1. The age for division was determined as the approximate median of the ages of children involved.

2.2. General procedures

Testing occurred either at a research laboratory or at the participants' preschool in a quiet corner of a room. Eight familiar objects served as well-known baseline stimuli. Participants reliably selected each baseline object when presented with the corresponding word. The baseline stimuli included a stuffed dinosaur, doll, toy airplane, spoon, glass, book, and toy car. Two nonsense words ("pafe" and "shede") served as the target words, and hard-to-label "junk objects" served as target referents (see Fig. 2).

Order of the trials and the set of stimuli to be presented on each trial were predetermined and were the same for all children within a condition. Four objects were presented in a line on the table in front of the child, and the child was asked to select one of them. The experimenter presented all target words within a consistent carrier phrase ("Can you show me the 'pafe'?") and recorded the child's responses. Children were not given feedback about the accuracy of their selections. Children chose a small prize at the end of the session.

Table 1
Participant characteristics of age-based (age in months) cohorts

Age cohort	Exposure condition	
	Concurrent	Successive
Older		
n	9	13
Mean age	47.1	49.6
Age range	26–40	29–41
Younger		
n	12	24
Mean age	35.1	35.1
Age range	43–52	43–57

Fig. 2. Photograph of sample stimuli used in the present study. In the foreground, second from the left and right, are the new novel objects, the "pafe" and "shede." In the back row are three of the familiar objects, a ball, cup, and toy airplane. Note: This is not how the objects were displayed during testing, where only four were presented at a time and all were presented in a line.

2.3. Concurrent versus successive introduction procedures

Each protocol (concurrent or successive introduction) was completed in a single session, with testing conducted directly after teaching. Three exposures were presented for each word, resulting in a total of six exposures altogether. All exposure trials were interspersed within baseline trials. At the end of the teaching block, four learning-outcome trials (two per word) were interspersed among three baseline trials.

Tables 2 and 3 illustrate the trials within each protocol (concurrent and successive, respectively) in the order these trials were presented. Descriptions of each trial and this structure are also provided in the text below.

2.4. Exposure trials: "Pafe"

The nonsense word "pafe" was the first nonsense word–object pairing participants encountered. The structure of the exposure trials for "pafe" was identical in both the concurrent and successive protocols; the novel target object was contrasted against three well-known, familiar objects. Distribution of the trials within the session differed, however. In the concurrent introduction protocol, exposures to the word "pafe" were intermixed throughout the teaching block, while in the successive introduction protocol these were the first three exposure trials the children received.

Table 2
Description of trial types and session structure: concurrent introduction

Trial type/order	Spoken word	Stimuli (probe trials only)			
		Choice #1	Choice #2	Choice #3	Choice #4
2 Baseline trials					
Exposure trial	“pafe”	Target #1	Toy car	Dinosaur	Toy plane
2 Baseline trials					
Exposure trial	“shede”	Target #2	Dinosaur	Toy plane	Spoon
1 Baseline trial					
Exposure trial	“shede”	Target #2	Ball	Book	Dolly
2 Baseline trials					
Exposure trial	“pafe”	Target #1	Spoon	Car	Cup
1 Baseline trial					
Exposure trial	“pafe”	Target #1	Cup	Dinosaur	Ball
2 Baseline trials					
Exposure trial	“shede”	Target #2	Ball	Book	Spoon
Learning outcome trial	“pafe”	Target #1	Target #2	Dinosaur	Doll
1 Baseline trial					
Learning outcome trial	“shede”	Target #2	Target #1	Ball	Spoon
1 Baseline trial					
Learning outcome trial	“pafe”	Target #1	Target #2	Book	Cup
1 Baseline trial					
Learning outcome trial	“shede”	Target #2	Target #1	Car	Toy plane

The stimuli listed in this table are the actual stimuli used on each trial; Correct choice is always listed in the first (left-hand) column, but during testing objects were arranged in a random order on the table.

2.5. Exposure trials: “Shede”

The nonsense word “shede” was the second nonsense word–object pairing participants encountered. In concurrent introduction, the trial structure for introducing “shede” was identical to that for the first word–object pairing (the “pafe” trials). That is, the novel target object was contrasted against three well-known, familiar objects (see Table 2). In successive introduction, however, the trial structure for introducing “shede” differed. Here, the novel target object was contrasted against the object that had just been labeled “pafe.” That is, on each of the “shede” exposure trials one of the other three objects in the choice array was the object previously identified as “pafe” (see Table 3). In addition, while in concurrent introduction, exposures to the word “shede” were intermixed throughout the teaching block; in successive introduction the three “shede” trials occurred after the three “pafe” trials.

2.6. Test trials

Testing for learning outcomes was identical irrespective of whether the words had been introduced through the concurrent or the successive introduction procedure. On test trials, both novel target objects appeared in the array, along with two familiar objects. Thus, when asked to identify the “pafe” or the “shede,” the child had to select between the two novel

Table 3
Description of trial types and session structure: Successive introduction

Trial type/order	Spoken word	Stimuli (probe trials only)			
		Choice #1	Choice #2	Choice #3	Choice #4
2 Baseline trials					
Exposure trial	"pafe"	Target #1	Toy car	Dinosaur	Toy plane
2 Baseline trials					
Exposure trial	"pafe"	Target #1	Cup	Dinosaur	Ball
1 Baseline trial					
Exposure trial	"pafe"	Target #1	Spoon	Car	Cup
2 Baseline trials					
Exposure trial	"shede"	Target #2	Target #1	Book	Dolly
1 Baseline trial					
Exposure trial	"shede"	Target #2	Target #1	Toy plane	Spoon
2 Baseline trials					
Exposure trial	"shede"	Target #2	Target #1	Ball	Spoon
Learning outcome trial	"pafe"	Target #1	Target #2	Dinosaur	Dolly
1 Baseline trial					
Learning outcome trial	"shede"	Target #2	Target #1	Ball	Spoon
1 Baseline trial					
Learning outcome trial	"pafe"	Target #1	Target #2	Book	Cup
1 Baseline trial					
Learning outcome trial	"shede"	Target #2	Target #1	Car	Toy plane

The stimuli listed in this table are the actual stimuli used on each trial; Correct choice is always listed in the first (left-hand) column, but during testing objects were arranged in a random order on the table.

targets on the basis of the spoken word. To select successfully, the child had to remember which object had been paired with which word. Four test trials (two per word) were presented. These trials were the first time that children in the concurrent introduction condition were presented with the two novel objects together.

2.7. Scoring and data analysis strategy

Data were analyzed separately for the participants' performance on the three exposure trials for each novel word as well as for the four test trials at the end of the session.

2.8. "Pafe" exposure trials

Performance on these three trials was analyzed as a continuous variable (0–3 trials correct) and as a binary variable (passed criterion or not). The binary outcome variable was adapted from Mervis and Bertrand (1994). They considered a child a "fast mapper" if the child selected correctly on three or four of four trials (in that study, each trial presented a new label/exemplar). In this study, correct selection on two or three of the three exposure trials was considered reliable performance. The number of children who succeeded and failed to meet this criterion was calculated for older and younger children in each condition.

For analysis of performance on "pafe" exposure trials as a continuous variable, Wilcoxon–Mann–Whitney was used. For analysis of performance here as a binary variable, logistic regression was used, with condition stratified within age and with an age × condition interaction term. For a 2 × 2 table, chi-square analysis is identical to logistic regression, thus chi-square was used in some instances. All analyses were done using exact inference methods, not asymptotic ones (StatXact; Mehta, 1994; Mehta & Patel, 1995, 2000). Because the trial structure for "pafe" trials was identical across the two conditions, performance here was expected to be comparable across the two experimental conditions.

2.9. "Shede" exposure trials

Performance on "shede" exposure trials was analyzed in the same way, using both the continuous and binary outcome measures described above for "pafe" exposure trials. Performance by children on "shede" trials in concurrent introduction should be similar to their performance on the "pafe" trials because in concurrent introduction the "shede" introduction was structured the same as "pafe." Of interest was children's performance on "shede" trials under the successive introduction protocol, as these required children to explicitly contrast the wholly novel "shede" object with the just-labeled "pafe" object. To compare performance within condition of the same children's performance on "pafe" and "shede" trials, the McNemar test for within-subject analyses was used.

2.10. Test trials

Trials testing learning outcome required participants to select alternately between the two novel objects on the basis of their labels. Performance on test trials was analyzed as a continuous outcome (0–4 trials correct) using Wilcoxon–Mann–Whitney, and as a binary variable (passed criterion or not) using logistic regression. A strict criterion for concluding that children had in fact remembered the new map was imposed; children were considered to have learned each new word-pair relation if they selected correctly on all four of the test trials. The number of children who succeeded and failed to meet this criterion was calculated for older and younger children in each condition.

3. Results

3.1. Performance on exposure trials: "Pafe"

3.1.1. Overall

Of the 21 children in concurrent introduction, 19 children (90%) met criterion for reliable performance on the "pafe" exposure trials, as did 33 of the 37 children in successive introduction (89%). Both Wilcoxon–Mann–Whitney and chi-square analyses confirmed that, as predicted, there was no difference in performance on these trials by children in the two conditions (ps > .10).

3.1.2. Age analysis

Fig. 3 illustrates the percentage of participants in each age group and in each condition who met the criterion for correct performance on exposure trials for "pafe" (two to three correct object choices). Of the older cohort of children (over 42 months), 100% met this criterion in each condition. Of the younger cohort of children (under 42 months), 83% selected correctly on at least two of the three "pafe" trials, under both the concurrent and successive conditions. Chi-square analysis indicated that the difference between the proportion of younger and older children who met criterion was not statistically significant ($p > .10$).

3.2. Performance on exposure trials: "Shede"

3.2.1. Overall

Of the 21 children in the concurrent introduction condition, 20 children (95%) met the criterion for reliable performance on "shede" exposure trials. Of the 37 children in the successive introduction group, 23 children (62%) met criterion on the "shede" trials within that protocol, which paired the label "shede" with a novel object, in the presence of the object previously called a "pafe."

Statistical analysis compared between-group performances on "shede" exposure trials across conditions. Performance was significantly better under the concurrent introduction condition than the successive introduction condition, both when performance was analyzed as a continuous variable (Wilcoxon–Mann–Whitney = 532.5, $p < .009$) or as a binary variable, χ^2 (1, N = 37) = 7.6; p = .006.

A within-group McNemar test was conducted to determine whether children's performances differed for "pafe" exposure trials compared to "shede" exposure trials. As expected, there were no significant differences between children's performances in concurrent introduction. The difference was significant for successive introduction (p = .002), with better performance on the "pafe" trials than the "shede" trials.

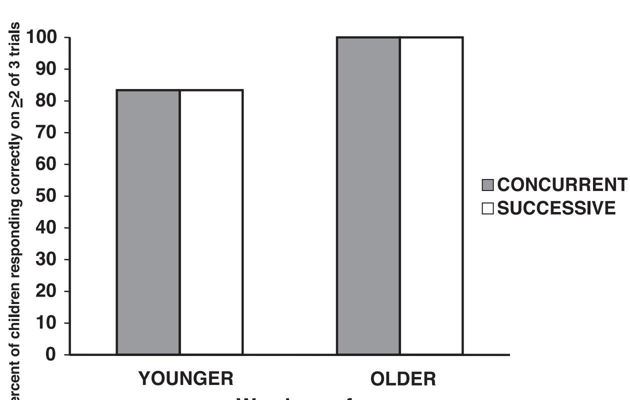

Fig. 3. Percentage of participants meeting criterion for reliable selection on the three trials introducing "pafe" and Target Object #1 by age and condition.

3.2.2. Age analysis

Fig. 4 illustrates the percentage of participants in each age group and each condition who selected correctly on at least two of the three exposure trials for "shede." All of the older children (100%) and 92% of the younger children selected correctly on at least two of the three "shede" exposure trials in the concurrent condition; this difference was not statistically significant. As predicted, these proportions are comparable to those reported for "pafe" trials (cf. Fig. 3). This was expected because in concurrent introduction the trial structure for introducing "shede" was identical to that for introducing "pafe."

However, when performance on "shede" trials in the successive introduction condition was evaluated, significant age-related differences emerged. While 85% of older children met criterion for reliable performance on "shede" exposure trials in successive introduction, only 50% of the younger children did so. Both Wilcoxon–Mann–Whitney and chi-square analysis confirmed that performance on "shede" exposure trials in successive introduction was better in older than younger children, Wilcoxon–Mann–Whitney = 272, $p = .03$ and χ^2 $(1, N = 37) = 3.7$; $p < .05$, and that performance did not differ significantly by age in the concurrent condition. Thus, the challenges of the "shede" trials in this protocol appeared to be particularly difficult for younger children, but not for older ones.

Overall, older children performed better (95% correct) than younger children (62% correct) on the "shede" exposure trials. Although there were a greater number of younger children in the successive than the concurrent condition, the effect of condition cannot be accounted for by age. Performance on "shede" exposure trials was better in concurrent than successive introduction even controlling for age (Wilcoxon–Mann–Whitney = 532.5, $p < .01$). When performance is judged by whether children passed criterion on "shede" exposure trials, again children in the concurrent condition outperformed children in the successive condition even controlling for age (Wald statistic = 4.6, $p < .03$, $df = 1$). The age by condition interaction was not significant, and using the binary outcome measure the difference in performance by age was no longer significant.

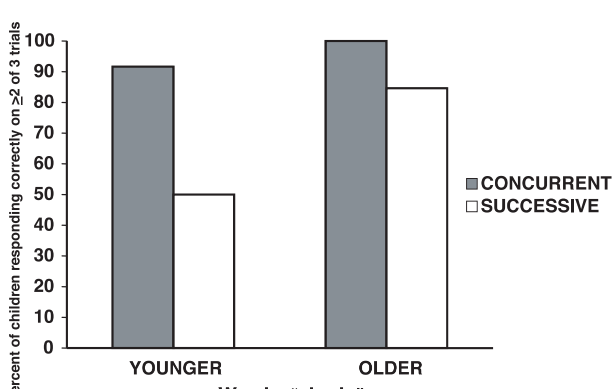

Fig. 4. Percentage of participants meeting criterion for reliable selection on the three trials introducing "shede" and Target Object #2 by age and condition.

Older children performed equally well on the "pafe" and "shede" trials overall and within each of the two conditions. Younger children, however, performed significantly better on "pafe" than on "shede" exposure trials (McNemar, $p < .001$). When analyzed by condition, the difference in younger children's performance was not significant for the concurrent condition, but it was significant for the successive condition ($p < .01$).

The overall group difference reported above on "shede" trials is accounted for by differences in the younger children's performance. When only the results for the older children are analyzed, the difference between children's performance on "shede" trials in the two conditions is not significant. That difference is significant when only the performance of younger children is analyzed, Wald statistic ($df = 1$) = 9.64, $p = .002$, Wilcoxon–Mann–Whitney = 84.0, $p < .05$.

3.3. Pattern of errors

In the successive condition, the four children who failed to meet criterion on "pafe" exposure trials also failed to meet criterion on "shede" exposure trials; the remaining 10 children who failed to meet criterion on "shede" had previously met criterion on the earlier "pafe" trials. In the concurrent condition, 3 children failed to meet criterion; 2 failed "pafe" only and 1 failed "shede" only. No child in the concurrent condition failed to reach criterion on both "pafe" and "shede" exposure trials. The mean age of the children who failed to meet criterion was 36 months (*SD* = 8 months) for both concurrent and successive introduction; all but two were in the "younger" cohort. Thus, the children who failed to pass criterion on the exposure trials were at the youngest end of the ages tested.

3.4. Performance on test trials

Tests for learning outcomes required children to alternately select between Target #1 and Target #2 objects on the basis of the spoken words "pafe" and "shede." It makes little sense to test for learning after exposure for children who did not select correctly on the exposure trials. Thus, learning outcome data are reported for children who met the criterion for mapping for both new words (two or three correct selections) separately from those who did not.

3.5. Children who met criterion on exposure trials

Forty-one of the 58 children in the sample (70%) met criterion for reliable fast mapping of both novel words during exposure. Table 4 presents the ages for this subset of children.

Test trial performance by these children was evaluated based on the number of correct selections out of a possible four opportunities. Fig. 5 illustrates the selection pattern for each individual child in each of the two exposure conditions. The *x*-axis represents chronological age of the individual child (in months). The number of correct selections is graphed along the *y*-axis. Each child is represented as a single data point, plotted by age and number of correct selections. The letters C and S represent which procedure the child in question underwent, concurrent or successive, respectively.

Table 4
Number and age (in months) of the children who met the criterion for reliable fast mapping in exposure trials

Group	Exposure condition	
	Concurrent	Successive
Total group		
n	18	23
Mean age	40.9	42.8
Age range	26–52	32–57
Age groups		
Older *n*	9	11
Mean age	47.1	49.5
Age range	43–52	43–57
Younger *n*	9	12
Mean age	34.7	36.6
Age range	26–40	32–40

For analysis of performance as a binary variable, criterion for passing testing for learning outcomes had been set at 100% correct (4/4). As can be seen in Fig. 5, more children overall met this criterion after successive introduction than after concurrent introduction, with only 28% of participants meeting criterion in the concurrent introduction but 56% of participants doing so after successive introduction. The difference between conditions was significant when performance on test trials was judged by whether criterion was met or not, Wald statistic ($df = 1$) = 3.98, $p = .046$, and that was true even controlling for age. The odds ratio produced by the logistic regression analysis indicated that children in the successive introduction condition were 11 times more likely to pass the criterion for learning on the test trials than children in concurrent introduction. (The difference between conditions just failed to reach statistical significance when performance on test trials was treated as a continuous measure, Wilcoxon–Mann–Whitney = 142.5, $p = .075$.)

When evaluated by age group, significant differences emerged for performance by children in the younger age group under the two exposure conditions. Fig. 6 presents the proportion of children in each group that met the criterion for selecting four of four trials correctly during the test trials. Older children performed equally well on learning outcome test trials following either type of exposure protocol, with 44% (concurrent) and 55% (successive) of the children in the older group meeting criterion. A greater number of younger children, however, met criterion after successive introduction (58% of children) than after concurrent introduction (only 11% of children). Logistic regression confirmed that the difference in performance across conditions was not significant for older children, but was for younger children, Wald statistic ($df = 1$) = 3.9; $p = .046$.[1] Younger children's fast mapping performance was better

[1] A similar difference was observed when a slightly more relaxed criterion of three of four correct selections was imposed. Seventy-three percent of older children met this criterion in successive introduction and 78% in concurrent introduction, while 83% of younger children met it in successive introduction but only 44% in concurrent introduction. Although the difference seems impressive, it just failed to meet statistical significance, χ^2 (1, $N = 21$) = 3.5; $p = .06$.

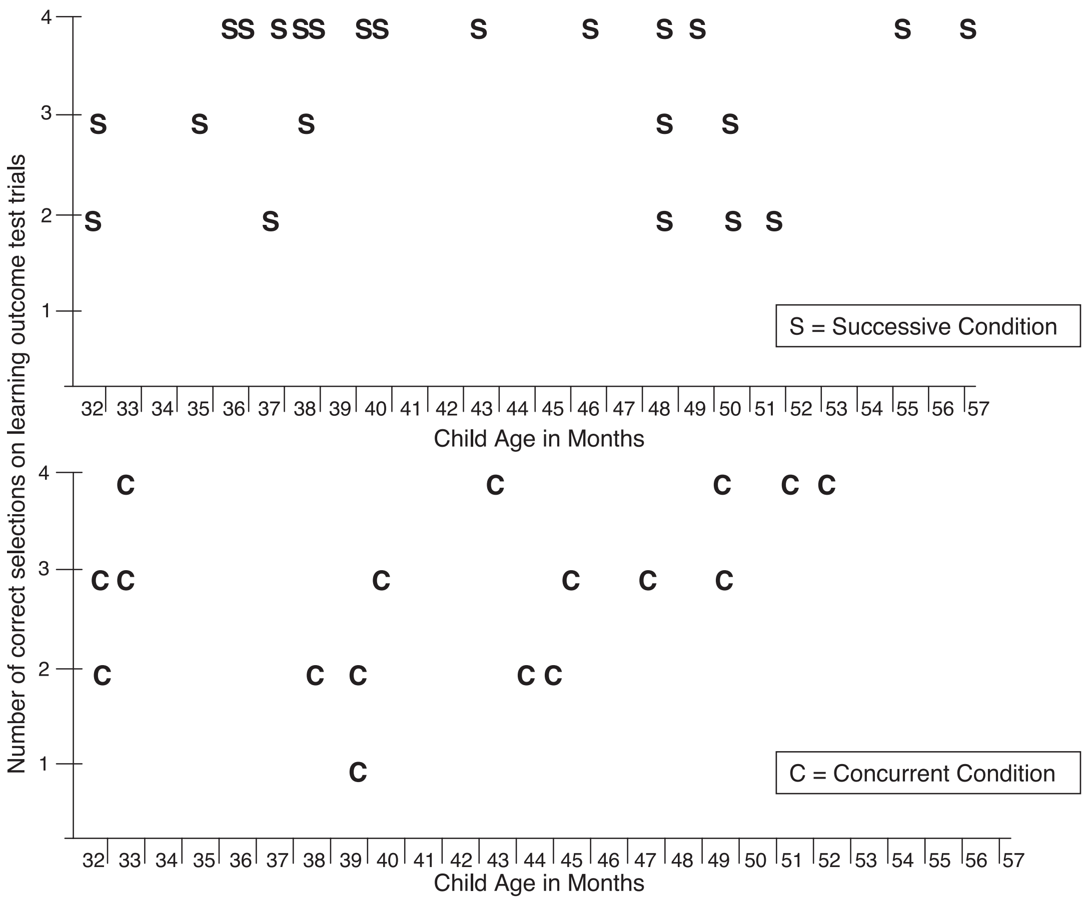

Fig. 5. Individual data for all participants on the test trials of learning outcome. Top panel presents the results for children in the successive condition. Bottom panel presents the results for those in the concurrent condition.

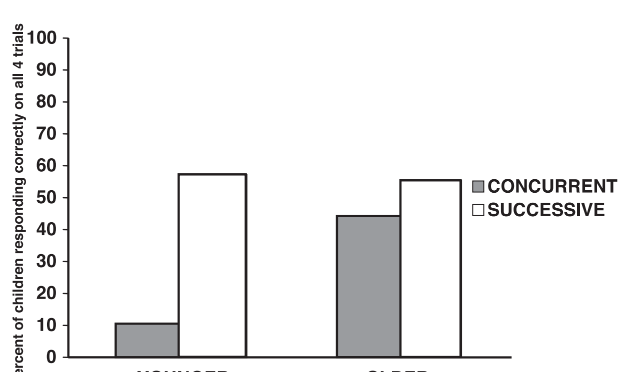

Fig. 6. Percentage of participants meeting the criterion for learning (100% correct) on the test trials by age and condition.

when they were exposed to the new words in the successive condition than under the concurrent exposure protocol. This effect held up for younger, but again not for older, children when test trial performance was treated as a continuous measure of learning (Wilcoxon–Mann–Whitney = 23.0, $p < .03$).

Overall, older children did not perform significantly better than younger children on the test trials, using either the continuous or binary outcome measure. Learning of children in the older group also did not differ significantly from that of younger children when only considering children in the successive condition. For children who had received the concurrent introduction procedure, however, younger children tended to perform somewhat worse at test than did older children (continuous measure of success: Mann–Whitney = 21.5, $p = .11$; criterion-based measure of success: Wald statistic [$df = 1$] = 2.19, $p = .13$).

3.6. Children who did not meet criterion on exposure trials

Not unexpectedly, children who failed to select correctly on the introductory exposure trials performed poorly on the test trials. None of the three children in the concurrent exposure condition and none of the 14 children in the successive exposure condition who failed to meet criterion during exposure met the criterion of selecting correctly on all four of the test trials. Only two of the participants met a more relaxed criterion of selecting correctly on three of the four test trials; both participants had undergone the successive exposure protocol.

4. Discussion

Previous research has established that fast mapping, which emerges during the second year of life, relies at least in part on children's appreciation of the difference between novel and known stimuli (Carey & Bartlett, 1978), though others have demonstrated that other cues can be used as well. The present study extends our understanding of the phenomenon by demonstrating that during the preschool years, children become increasingly efficient at using the novel/known contrast to map more than one new word. Children of 3–4 years were better able to quickly map a second novel word and to retain the two new word–referent pairings regardless of the training procedure used than were younger children. Younger children were far more sensitive to the particular procedure for introducing the novel word–object pairings. The present study also extends our understanding of fast mapping by providing information about when rapid mappings of multiple words are more likely to occur and what kind of training procedure is more likely to be efficacious.

4.1. Fast mapping and the procedural differences

If we were to look only at exposure trials, we might conclude that concurrent introduction is superior, resulting in reliable selections for both target words. However, those reliable selections on exposure trials were often not sufficient to result in children reliably recalling which word had been paired with which target when tested on learning outcome test trials. A mere 11% of the younger children selected correctly on the testing trials after concurrent introduction, even though they had all correctly paired each new word with its target on at least four of the six exposure trials (at least two per word). In contrast, children in successive introduction who had selected correctly on the same minimal number of exposure trials were significantly more likely to then recall which word was paired with which target. In this case, 58% of children met criterion for reliable performance on learning outcome, a proportion similar to that for older children.

The differences in performance of children in the two procedures likely reflects the change in trial structure by which the "shede" (second novel target) was introduced. Recall that the trial structure for "pafe" exposure contrasted the novel target (Target #1) with only highly familiar objects like ball and plane. For children who are able to make use of a contrast between new and known labels and objects, the correct selection on such trials is obvious. For preschool-aged children like those tested here, we expected and obtained high numbers of children demonstrating the use of this strategy.

The same logic holds true for the "shede" exposure trials in concurrent introduction, which were structured identically to "pafe," and again we found highly accurate performances on these trials. However, in successive introduction the task was a bit more difficult. The child was asked now to select the wholly novel target (Target #2) when one of the options available in the array was the object that had just been labeled "pafe." Only if children (a) distinguish the small difference in the relative novelty of the two targets and (b) remember more about Target #1 than that it was novel (e.g., remember features of its appearance) will they select correctly on such a trial. Our data indicated that although older children had little trouble with this, younger children found it challenging. Presence of the basic insight concerning novel/known items and labels (demonstrated in "pafe" trials) did not guarantee that when two relatively novel words were presented younger children could use the information that had been provided to deduce the correct word–target map.

Finally, the test trials presented the greatest challenge: Given that children had all demonstrated accurate selections during exposure trials, test trials examined whether they remembered which novel item was paired with which novel word. Once again, older children succeeded irrespective of how they had been introduced to the new words. Younger children, however, were far more likely to recall the word–target pairings if they had undergone successive introduction rather than concurrent introduction. Although more children of just barely 3 years failed to reach criterion for "reliable mapping" (two or three correct of three trials) during exposure trials for "shede" in the successive than concurrent introduction, among those children who did demonstrate reliable selection on exposure trials, the successive introduction procedure resulted in significantly better outcomes in the test trials for learning.

The data converge on the conclusion that although concurrent introduction offers children an immediate means to guess a new word's meaning, it is less likely to result in long-term gains in vocabulary, at least in younger children. It appears that this is because the concurrent introduction procedure requires that the children do no more than recognize that a target object is novel. Children can succeed perfectly well on concurrent exposure trials by paying the most minimal attention to the features of any novel object and without ever remembering which novel word was paired with which novel object. The successive introduction procedure, however, includes an additional trial type (the modified "shede" trials) that cue the child into the contrast to be tested later.

The basic "pafe" trials in this study revealed whether or not each child was making use of the contrast between known and novel words to help learn new words. Use of this contrast would be the first step in fast mapping, and is an important early strategy. The modified "shede" trials in the successive condition revealed the extent to which the three exposures to "pafe" in the novel/known contrast resulted in an actual "fast map" of the word "pafe" to the first target. This serves as a second step, in which fast mapping is applied for purposes of actual word learning. Finally, the test trials revealed how well the two new words were maintained after three exposures to each new word. The ability to select between two alternatives conditional upon the word spoken represents most strictly a test of whether or not the novel/known contrast resulted in word learning.

The present study illustrates that inclusion of the modified "shede" trials offers a view of how children progress from a simple appreciation of novel/known contrast through an ability to make full use of it for purposes of word learning. The concurrent introduction procedure, which does not incorporate this intermediate trial type, is less sensitive to this potential developmental step in the application of fast mapping for word learning. Thus, the successive procedure not only appears to yield a better learning outcome but provides a more sensitive measure of developmental or group differences because it provides information about an intermediate level of acquisition that is not possible to access in the concurrent condition.

4.2. Caveats and future research directions

Several methodological caveats are worth noting, as are potential research directions. These include the stimuli, number of exposure trials, and methodological comparison we used.

In our design we did not counterbalance the two objects that served as "pafe" and "shede." It is possible that one of the objects (the taller "pafe" object) might have been more salient to children than the other, accounting for more accurate selections on "pafe" than "shede" trials overall. However, that difference was found in only one condition. Only in the successive introduction condition did we see errors specific to "shede" and its target, primarily among younger children. It is possible that performance was artificially depressed on "shede" exposure trials in the successive condition (where the "pafe" object was also visible) because the object paired with the word "shede" was not sufficiently salient.

An important direction for future research concerns the number of exposures to use, in research and in teaching. Prior research (Wilkinson & Green, 1998) presented children with 24 exposures, which would seem excessive for typically developing children. However, the findings of the present study suggest that for many children, even those 3.5–4 years old (but especially for those just barely 3 years old), 3 exposures per word is not enough to reliably learn two new words. Rice and colleagues (1994) evaluated learning of multiple new words using differential numbers of exposures, but they were teaching many new words. Might 5 exposures to each new word be sufficient? Does this change with age? It would be of interest, particularly from an instructional standpoint, to examine the relationship between number of exposures and reliability of learning. It would also be a stronger test of whether successive introduction really does promote better learning outcome if children in both the concurrent and successive conditions performed well enough during exposure so that all or almost all of the children in both conditions could be included in analyses of learning outcome on the test trials.

Another important direction for future research might be to clue children in earlier that they will need to encode the association between a specific new word and a specific new object (rather than just relying on relative novelty). That is, instead of presenting the "pafe" trials and only after that first presenting two novel objects in the array for the first "shede" trial, it might help children to learn and remember the word–referent pairings if the first "shede" trial in the successive condition were administered before the second "pafe" trial. Performance would likely be poor on that "shede" trial, but presenting the trial early in the session might well lead to a better learning outcome.

4.3. Relationship of these findings to learning more generally

Although some have argued that fast mapping is a strategy specific to language learning (e.g., Markman, 1992), the pattern of results we have found certainly shares characteristics of learning more broadly. We are suggesting that a comparison set consisting of familiar objects plus a novel one requires the child to draw a different distinction than does a comparison set of only familiar objects. This is consistent with the observation that if information is not relevant for action, then a viewer is less likely to pay attention to it. The information needed for action is the information needed to choose among alternative possible responses. Hence, the set of alternatives (the comparison set) critically affects what one notices and learns (Olson, 1964). Certainly there is evidence in the literature on the facilitative effects of what have been called "contrastive learning procedures" for performance on cognitive and information-processing tasks (e.g., Soraci, Deckner, Baumeister, & Carlin, 1990).

The pattern of our results would suggest that although all the information about the features of the two nonsense objects was available to participants in the concurrent procedure, they did not pay attention to that information. The Diamond laboratory has been studying what cues and information children are able to benefit from at different ages and in different circumstances, and what strategies children at different ages adopt in solving various problems posed by a set of tasks. In one such task, a sticker is placed in each of six boxes. The child's job is to find all the stickers in the least number of reaches, that is, without reaching to any box more than once. In one variant of the task, after the child watches the experimenter fill and close each box, the experimenter puts an "apple" sticker on the lid of each box. During the course of testing, when the child finds and removes the sticker from any box, the experimenter removes the apple sticker from the top of that box and replaces it with a sticker depicting an apple core. This is meant to remove the memory demands of the task by giving children a visible cue of which boxes are empty (those in which the "apple" had been "eaten") and which are still "full" with a sticker inside. Children of 2 1/2–3 1/2 years, however, perform no better with the visible cues than without them. They do not appear to notice or use the available information; instead they appear oblivious to it (Jacobs, Fahey, Groff, Stollstorff, & Diamond, 2003).

One small change makes all the difference in the world, just as does our one small change in our fast-mapping procedures. The change for the six-box task is simply to tell the children explicitly what the "full-apple" and "apple-core" stickers are meant to represent. Indeed, preliminary results indicate that children of 2–3 years are able to make full use of utterly arbitrary symbols for "full" and "empty" (red and blue stickers), *if* the children are explicitly told what they represent (Jacobs et al., 2003). In short, when children's attention is not explicitly drawn to the cues, children of 2–3 years do not appear to show a benefit from them in their performance. When their attention is explicitly drawn to the cues, and their symbolic meaning explicitly explained, children of 2–3 years are able to remember what the cues mean, attend to them, and use them to guide their performance.

Also, our results are similar to findings in other aspects of learning more generally in that the condition that produces better immediate performance is often not the condition that produces the better long-term outcome. "Shede" exposure trials in the successive condition were more difficult and younger children performed worse on them. However, those harder trials produced a better learning outcome on test trials. Similarly, random or spaced practice produces worse immediate performance than does blocked or massed practice, but the former leads to better long-term retention (in verbal learning: Landauer & Bjork, 1978; Rea & Modigliani, 1985, and motor learning: Magill & Hall, 1990; Shea & Morgan, 1979). Providing frequent feedback results in good short-term retention but poor long-term retention; providing infrequent feedback leads to good long-term retention but poor short-term retention (on cognitive tasks: Krumboltz & Weisman, 1962; Schooler & Anderson, 1990; Schulz & Runquist, 1960; on motor tasks: Schmidt, Young, Swinnen, & Shapiro, 1989; Winstein & Schmidt, 1990; Wulf & Schmidt, 1989). Using the same materials (constant training) is easier and leads to better performance just after training but worse performance in later tests. Variable training is harder and results in worse immediate performance but better performance in later tests (in the cognitive domain: Bransford, Franks, Morris, & Stein, 1977; in the motor domain: Shapiro & Schmidt, 1982).

4.4. Implications for teaching children

Teachers and parents should not assume that just because information is perceptually available that a child will notice it and/or remember it. The children in our concurrent condition had as much information perceptually available to them as children in the successive condition but the younger children in the concurrent condition did not pick up on or retain that information, whereas more of the younger children in the successive condition did. If we as instructors want a child to notice and remember something (be it the appearance of something, its function, its spatial location, or anything) then we must make that something relevant to the choices the child must make, relevant to the actions the child must take. The study by Jacobs et al. (2003) indicates that sometimes even that is not sufficient. Sometimes a child needs us to point out that, or in what way, the information that is available is relevant to what the child is doing. Sometimes children do not realize that relation on their own. As we learn more about what helps children notice such relations, we will be progressively better able to provide children with the structure they need in order to learn.

Acknowledgements

Data collection was supported by NICHD R01 HD35453 (to AD). Preparation of the manuscript was supported by NICHD R01 HD35453 and by NICHD HD R21 35015 (to KMW). We would like to thank Kristin Shutts for help with data collection, coding, and data entry, and all the parents and children whose cooperation made this research possible.

References

Bates, E., Dale, P., & Thal, D. (1995). Individual differences and their implications for theories of language development. In P. Fletcher, & B. MacWhinney (Eds.), *The handbook of child language* (pp. 96–151). Cambridge, MA: Blackwell.

Bransford, J. D., Franks, J. J., Morris, C. D., & Stein, B. S. (1977). Some general constraints on learning and memory research. In L. S. Cermak, & F. I. M. Craik (Eds.), *Levels of processing in human memory* (pp. 331–354). Hillsdale, NJ: Erlbaum.

Carey, S. (1982). Semantic development: The state of the art. In E. Wanner, & L. R. Gleitman (Eds.), *Language acquisition: The state of the art* (pp. 265–293). Cambridge, MA: MIT Press.

Carey, S., & Bartlett, E. (1978). Acquiring a single new word. *Papers and Reports on Child Language Development*, *15*, 17–29.

Dickinson, D. K. (1984). First impressions: Children's knowledge of words gained from a single exposure. *Applied Psycholinguistics*, *5*, 359–373.

Dollaghan, C. (1985). Child meets word: "Fast mapping" in preschool children. *Journal of Speech and Hearing Research*, *28*, 449–454.

Gibson, E. J., & Pick, A. D. (2000). *An ecological approach to perceptual learning and development*. London: Oxford University Press.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.

Golinkoff, R. M., Hirsh-Pasek, K., Bailey, L. M., & Wenger, N. R. (1992). Young children and adults use lexical principles to learn new nouns. *Developmental Psychology*, *28*, 99–108.

Golinkoff, R. M., Mervis, C. B., & Hirsh-Pasek, K. (1994). Early object labels: The case for a developmental lexical principles framework. *Journal of Child Language*, *21*, 125–155.

Hollich, G. J., Hirsh-Pasek, K., Golinkoff, R. M., Brand, R. J., Brown, E., Chung, H. L., Hennon, E., & Rocroi, C. (2000). Breaking the language barrier: An emergentist coalition model for the origins of word learning. *Monographs of the Society for Research in Child Development*, *65*(5, Whole).

Jacobs, E., Fahey, M., Stollstorf, M., & Diamond, A. (2003). *Young children's inferential and observational abilities as revealed by performance on a six-box self-ordered retrieval task*, Manuscript in preparation.

Kagan, J. (1981). *The second year: The emergence of self-awareness*. Cambridge, MA: Harvard University Press.

Krumboltz, J. D., & Weisman, R. G. (1962). The effect of intermittent confirmation in programmed instruction. *Journal of Educational Psychology*, *53*, 250–253.

Landauer, T. K., & Bjork, R. A. (1978). Optimal rehearsal patterns and name learning. In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory* (pp. 625–632). London: Academic Press.

Lederberg, A. R., Prezbindowsi, A. K., & Spencer, P. E. (2000). Word learning skills of deaf preschoolers: The development of novel mapping and rapid word learning strategies. *Child Development*, *71*, 1571–1585.

Magill, R. A., & Hall, K. G. (1990). A review of the contextual interference effect in motor skill acquisition. *Human Movement Science*, *9*, 241–289.

Markman, E. M. (1989). *Categorization and naming in children*. Cambridge, MA: MIT Press.

Markman, E. M. (1992). Constraints on word learning: Speculations about their nature, origins, and domain-specificity. In M. R. Gunner, & M. P. Maratsos (Eds.), *Minnesota symposium on child psychology* (pp. 59–101). Hillsdale, NJ: Erlbaum.

Mehta, C. R. (1994). The exact analysis of contingency tables in medical research. *Statistical Methods for Medical Research*, *3*, 135–156.

Mehta, C. R., & Patel, N. R. (1995). Exact logistic regression: Theory and examples. *Statistical Medicine*, *14*, 2143–2160.

Mehta, C. R., & Patel, N. R. (2000). *StatXact 4 for Windows: Statistical software for exact nonparametric inference user manual*. Cambridge, MA: CYTEL Software.

Mervis, C. B., & Bertrand, J. (1994). Acquisition of the novel-name–nameless-category (N3C) principle. *Child Development*, *63*, 1646–1662.

Mervis, C. B., & Bertrand, J. (1995). Acquisition of the novel-name–nameless-category (N3C) principle by young children who have Down syndrome. *American Journal on Mental Retardation*, *100*, 231–243.

Olson, D. R. (1964). *Cognitive development: The child's acquisition of diagonality*. New York: Academic Press.

Rea, C. P., & Modigliani, V. (1985). The effect of expanded versus massed practice on the retention of multiplication facts and spelling lists. *Human Learning: Journal of Practical Research and Applications*, *4*, 11–18.

Rice, M. (1989). Children's language acquisition. *American Psychologist*, *44*, 149–156.

Rice, M. (1990). Preschooler's QUIL: Quick incidental learning of words. In G. Conti-Ramsden, & C. E. Snow (Eds.), *Children's language* (*vol. 7*, pp. 171–195). Hillsdale, NJ: Erlbaum.

Rice, M. L., Buhr, J. C., & Nemeth, M. (1990). Fast mapping word-learning abilities of language-delayed preschoolers. *Journal of Speech and Hearing Disorders*, *55*, 33–42.

Rice, M. L., Oetting, J. B., Marquis, J., Bode, J., & Pae, S. (1994). Frequency of input effects on word comprehension of children with specific language impairment. *Journal of Speech and Hearing Research*, *37*, 106–122.

Schmidt, R. A., Young, D. E., Swinnen, W., & Shapiro, D. C. (1989). Summary knowledge of results for skill acquisition: Support for the guidance hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*, 352–359.

Schooler, L. J., & Anderson, J. R. (1990). The disruptive potential of immediate feedback. *Proceedings of the Cognitive Sciences Society* (pp. 702–708). Hillsdale, NJ: Erlbaum.

Schulz, R. W., & Runquist, W. N. (1960). Learning and retention of paired adjectives as a function of percentage occurrence of response members. *Journal of Experimental Psychology*, *59*, 409–413.

Shapiro, D. C., & Schmidt, R. A. (1982). The schema theory: Recent evidence and developmental implications. In J. A. Kelso, & J. E. Clark (Eds.), *The development of movement control and coordination* (pp. 113–150). New York: Wiley.

Shea, J. B., & Morgan, R. L. (1979). Contextual interference effects on the acquisition, retention, and transfer of a motor skill. *Journal of Experimental Psychology: Learning and Memory*, *5*, 179–187.

Soraci, S. A., Deckner, C. W., Baumeister, A. A., & Carlin, M. T. (1990). Attentional functioning and relational learning. *American Journal on Mental Retardation*, *95*, 304–314.

Wilkinson, K. M., & Albert, A. (2001). Adaptations of "fast mapping" for vocabulary intervention with augmented language users. *Augmentative and Alternative Communication*, *17*, 120–132.

Wilkinson, K. M., & Green, G. (1998). Implications of "fast mapping" for vocabulary expansion in individuals with mental retardation. *Augmentative and Alternative Communication*, *14*, 162–170.

Wilkinson, K. M., & Shah, N. (2003). *Mutual exclusivity and fast mapping in monolingual, bilingual, or second-language learning children*, Manuscript submitted for publication.

Winstein, C. J., & Schmidt, R. A. (1990). Reduced knowledge of results enhances motor skill learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*, 667–691.

Wulf, G., & Schmidt, R. A. (1989). The learning of generalized motor programs: Reducing the relative frequency of knowledge of results enhances memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*, 748–757.