Use of the mutual exclusivity assumption by young word learners

Ellen M. Markman, * Judith L. Wasow, and Mikkel B. Hansen

Department of Psychology, Stanford University, Jordan Hall, Bldg. 420, Main Quad, Stanford, CA 94305-2130, USA

Accepted 6 January 2003

Abstract

A critical question about early word learning is whether word learning constraints such as mutual exclusivity exist and foster early language acquisition. It is well established that children will map a novel label to a novel rather than a familiar object. Evidence for the role of mutual exclusivity in such indirect word learning has been questioned because: (1) it comes mostly from 2 and 3-year-olds and (2) the findings might be accounted for, not by children avoiding second labels, but by the novel object which creates a lexical gap children are motivated to fill. Three studies addressed these concerns by having only a familiar object visible. Fifteen to seventeen and 18–20-month-olds were selected to straddle the vocabulary spurt. In Study 1, babies saw a familiar object and an opaque bucket as a location to search. Study 2 handed babies the familiar object to play with. Study 3 eliminated an obvious location to search. On the whole, babies at both ages resisted second labels for objects and, with some qualifications, tended to search for a better referent for the novel label. Thus mutual exclusivity is in place before the onset of the naming explosion. The findings demonstrate that lexical constraints enable babies to learn words even under non-optimal conditions—when speakers are not clear and referents are not visible. The results are discussed in relation to an alternative social-pragmatic account.

© 2003 Elsevier Science (USA). All rights reserved.
1. Introduction

One of the major accomplishments of early childhood is the acquisition of language including the learning of thousands of words (Anglin, 1993). Once children become reasonably proficient speakers, there are several sources of information they can use to narrow down the meaning of a novel term. For children who have acquired enough syntax, grammatical form class cues can limit the hypotheses they consider, with count nouns more likely to refer to objects, adjectives more likely to refer to properties, and verbs more likely to refer to actions (Brown, 1957). In fact, distinctions within syntactic frames lead to more refined hypotheses about a verb’s likely meaning (Fisher, Gleitman, & Gleitman, 1991; Fisher, Hall, Rakowitz, & Gleitman, 1994). For older children, morphology becomes an important means of expanding their lexicon (Anglin, 1993). Contrast within a lexical domain (e.g., “that’s not red it’s maroon”) could provide further specification of a term’s meaning, narrowing it down from any property, to a more specific domain, say color (Au, 1990; Au & Laframboise, 1990; Carey & Bartlett, 1978; Heibeck & Markman, 1987). Social-pragmatic cues which reveal information about the communicative intent of the speaker are another source of information that, along with the non-linguistic context, can elevate some hypotheses about a word’s meaning over others (Baldwin, 1995; Tomasello & Akhtar, 1995). Given the importance of word learning, it makes sense that there would be multiple and often redundant sources of information that children can use to interpret a novel term (Markman, 1992; Woodward & Markman, 1998). But consider very young children who may find themselves in situations where these sources of information are uninformative, ambiguous, or even non-existent. To account for the success of these young children, several investigators have postulated that children rely on some constraints on word meanings to shape their initial hypotheses about the meaning of novel words (e.g., Bloom, 1994; Clark, 1987, 1988; Markman, 1992, 1994; Markman & Hutchinson, 1984; Markman & Wachtel, 1988; Mervis, 1987; Waxman & Gelman, 1986; Waxman & Kosowski, 1990). These word learning biases serve as default assumptions or good first guesses as to a word’s meaning (Merriman & Bowman, 1989; Woodward & Markman, 1991). They help the child solve this inductive problem by narrowing the hypothesis space which, especially in absence of other sources of information about a word’s meaning, would be overwhelmingly large.

The focus of this paper is on one of the postulated constraints: mutual exclusivity, which leads children to prefer that each object have only one category label (Au & Glusman, 1990; Liittschwager & Markman, 1994; Markman, 1989, 1992; Merriman & Bowman, 1989; Merriman & Schuster, 1991). Mutual exclusivity enables children to avoid redundant hypotheses (Liittschwager & Markman, 1994), helps children narrow overextension of words (Clark, 1983, 1987; Merriman & Bowman, 1989), and allows children to overcome the whole-object assumption thus freeing them to acquire words for parts, substances, and other properties (Markman & Wachtel, 1988; Soja, Carey, & Spelke, 1991; Woodward, 1992).

Another function of mutual exclusivity and other lexical constraints is that they enable children to successfully infer the referents of novel terms even under less than
optimal conditions such as the absence of direct referential cues from the speaker. That is, there are circumstances in which children can determine what object a novel word maps onto without the speaker pointing to or otherwise directing the child’s attention to the object. Suppose, for example, a child sees two objects, one of which is familiar, say a ball, and another novel, say a whisk, and hears someone say “Can you hand me the whisk?” According to mutual exclusivity a child should reject a second label for ball and thus infer that “whisk” might refer to the whisk given it is the only other object around. In fact there are now several powerful demonstrations that young children will treat a novel term as referring to the novel rather than the familiar object (Golinkoff, Hirsh-Pasek, Bailey, & Wenger, 1992; Markman & Wachtel, 1988; Merriman & Bowman, 1989; Merriman & Schuster, 1991). The role mutual exclusivity plays in such indirect word learning is the main focus of this paper.

Despite the evidence for the various roles that mutual exclusivity plays in young children’s word learning, a number of concerns have been raised about the argument that mutual exclusivity serves as a fundamental constraint on word learning. One widespread objection is that there are exceptions to this principle found in children’s vocabularies (Banigan & Mervis, 1988; Gathercole, 1987, 1989; Mervis, 1987; Mervis, Golinkoff, & Bertrand, 1994a; Nelson, 1988). Violations of the constraint found in a child’s lexicon, however, are not necessarily evidence against the existence of the constraint. To claim that children are biased to treat object labels as mutually exclusive is not to claim that they can never learn more than one label for the same object. To determine whether a child uses mutual exclusivity requires examining the initial hypotheses children consider. Children may well ultimately be successful at overriding the mutual exclusivity assumption yet nevertheless rely on mutual exclusivity as an initial guide to interpreting a novel label (see Woodward & Markman, 1991, and Merriman & Bowman, 1989, for elaborations of this point).

A second objection to the earlier work on mutual exclusivity is that it was consistently demonstrated only for children over 2 1/2 years old yet children become capable of rapid acquisition of words by roughly 18–19 months of age (Bloom, Lifter, & Broughton, 1985; Corrigan, 1983; Dromi, 1987; Goldfield & Reznick, 1990; Halliday, 1975; McShane, 1979; Nelson, 1973). Paralleling the doubt that young children use mutual exclusivity, is the claim that mutual exclusivity is simply a learned heuristic that older children might find useful, rather than a constraint that is essential for early word learning (MacWhinney, 1989). If mutual exclusivity is to play a role in the early rapid acquisition of vocabulary, it needs to be available to children well before 2 1/2 years. Several investigators have expressed skepticism that mutual exclusivity is available to younger children (Banigan & Mervis, 1988; Evey & Merriman, 1998; MacWhinney, 1989; Merriman & Bowman, 1989; Mervis & Bertrand, 1994; Mervis et al., 1994a; Mervis, Johnson, & Mervis, 1994b; Nelson, 1988), basing their doubt in part on diary data and experimental findings that children are capable of acquiring second labels for things.

The issues of second label learning and the age at which mutual exclusivity first occurs were both addressed in a series of studies by Liittschwager and Markman (1994). The studies were designed to monitor children’s acquisition of new object labels at the time of the first exposure to the novel words. As a test of whether
children prefer words to be mutually exclusive, children’s success at learning second labels for objects was compared to their success at learning first labels. In the first study, 24-month-olds were taught a novel term either for an object that they could already name or for one that they could not already name. Contrary to the predictions from mutual exclusivity, these 24-month-olds learned second labels as well as they learned first labels. These findings were surprising not just because they went counter to the predictions but because they failed to replicate the Banigan and Mervis (1988) finding that simple labeling of an object with a known first label was not sufficient for children to acquire the new label. Liittschwager and Markman explain these unexpected results by suggesting that depending on the information processing requirements of the task, 24-month-olds may be capable of overriding the mutual exclusivity assumption. In Banigan and Mervis’ (1988) study, children were taught several new terms, while in Liittschwager and Markman’s (1994) study, they were taught only one. This hypothesis was tested in a second study in which 24-month-olds were taught two rather than one new word—either two new first labels for two different objects or two new second labels for two different objects. Increasing the information processing demands of the task should make it harder for children to override mutual exclusivity and thus harder for them to learn second compared to first labels. This was exactly what was found. The 24-month-olds readily learned two new first labels for two different objects, but in marked contrast failed to acquire the second labels. Building on these findings, Liittschwager and Markman reasoned that younger babies with more limited information processing abilities should be less able to override mutual exclusivity even when acquiring a single new word. They tested this in a third study where 16-month-olds were taught a single new word—either a first label for an object or a second label for an object whose name they already knew. As predicted, 16-month-olds successfully acquired first but not second labels.

Thus, mutual exclusivity is used by babies as young as 16 months of age to reject second labels for things and thus avoid redundant hypotheses. It remains an open question, however, whether babies this young can benefit from the other functions of mutual exclusivity. One goal of the present work is to address this question with respect to indirect word learning.

Golinkoff, Mervis, and Hirsh-Pasek (1994) argue that such indirect word learning will not be achieved by children before the time of the vocabulary spurt. Their framework postulates a novel name-nameless category (N3C) principle, which holds that novel labels map on to novel objects and makes similar predictions as mutual exclusivity in many situations. On their account N3C is a second-tier principle that becomes available to children at the time of the vocabulary spurt and not as part of the first-tier principles available to children at the very beginning of language learning. To test this, Mervis and Bertrand (1994) presented 16–20-month-olds with displays of four familiar objects and one unfamiliar object and provided a novel label to see if they would map the novel label to the unfamiliar object. Following each task the experimenter provided the children with feedback by directly labeling the unfamiliar referent, this task becomes a direct teaching task.
Their results were that children who correctly mapped the label showed evidence of having entered the vocabulary spurt, producing an average number of 95 words according to parental reports. Non-mappers produced an average of 45 words. According to Mervis and Bertrand these results indicate that the ability to use N3C in indirect word learning becomes available only at the time of or after the child has entered the vocabulary spurt. However, this procedure likely yields a conservative estimate of children’s ability to disambiguate novel labels. As noted by Evey and Merriman (1998) and Graham, Poulin-Dubois, and Baker (1998), children may fail at this fast mapping task for a variety of reasons such as lack of processing capacity to manipulate four familiar and one unfamiliar objects. This explanation gains further credibility from another of Mervis and Bertrand’s findings. After the first exposure to the items where children had to infer the referent, children were given ostensive cues and feedback as to which novel object the label referred. Nevertheless, children still failed to learn. On Mervis and Bertrand’s logic then, we would have to conclude that children cannot learn words at all prior to the vocabulary spurt. Given the explicit teaching of the words, not only the fast-mappers but all children should have learned that the novel words refer to the novel objects.

To reduce the information processing load, Graham et al. (1998) presented children in the same age range with an indirect word learning task with three objects rather than five. They too found an effect of vocabulary size, but in contrast to Mervis and Bertrand’s findings, non-mappers had a mean productive vocabulary of 82 words and fast-mappers a mean productive vocabulary of 237 words. But Graham et al.’s finding that non-mappers produce an average of 82 words is not compatible with the generally accepted cut-off of 50 words for the vocabulary spurt. Furthermore, a vocabulary of 82 words in Graham and colleagues’ non-mappers is very close to the vocabulary of 95 words in children considered fast-mappers in Mervis and Bertrand’s study. Taken together the studies of Mervis and Bertrand and Graham et al. do not provide a coherent picture of children’s success at indirect word learning at different levels of vocabulary development and varying levels of task difficulty. The current studies attempt to resolve this issue.

Again, the earlier results are overwhelmingly that from about 2 1/2 years on, children consistently treat the novel word as referring to the novel object on indirect word learning tasks. Although this is consistent with the predictions from mutual exclusivity, there are three other explanations for children’s performance on this particular task.

Merriman and Bowman (1989) provided an important alternative and very different explanation for these results. They point out that these results could be obtained if children were predisposed to fill lexical gaps. One critical difference between the two hypotheses is that in this context mutual exclusivity and other lexical constraints provide a backup strategy for children to help them learn a new word even when no obvious candidate referent is around. In marked contrast, the lexical gap hypothesis is dependent on the presence of a candidate object to label. The lexical gap hypothesis concerning the indirect word learning situation is derived from a more general principle that states that in the presence of an object that as yet has no known label, children are motivated to discover its name (cf. Clark, 1983, 1987). The assumption
that things have names is related to the development of the “nominal insight” that McShane (1979) and others have postulated. Many children early in the course of language acquisition actively request labels for things. Thus, in the indirect word learning tests of mutual exclusivity children could map the novel word to the novel object because they have a novel object that they want to find a name for. If children had no reluctance to have second labels for things but simply were motivated to find first labels for things, then the results of these studies would be the same. Thus the lexical gap hypothesis might explain the results of the indirect word learning studies that have been interpreted as evidence in favor of mutual exclusivity. Whether children are in fact using mutual exclusivity in these cases is an open question.

One main goal of the present studies was to test whether the lexical gap hypothesis is a viable alternative explanation for children’s success on indirect word learning tasks. To distinguish between mutual exclusivity and the lexical gap hypothesis, the indirect word learning procedure was modified such that babies heard a novel label in the presence of a familiar object but no novel object was visible. If younger children are guided by mutual exclusivity to reject second labels for objects then they should avoid selecting the visible familiar object in the presence of a novel label. The lexical gap explanation is that upon seeing a novel object, children are motivated to find out what it is called. It could not explain a rejection of second labels in this situation because no novel object is visible to create a gap in the child’s lexicon.

Mutual exclusivity works in conjunction with the whole-object and the taxonomic assumptions, which state that words refer to whole objects of like kind. Implicit in these assumptions is the notion that children take words to refer. In the lexical principles framework (e.g., Golinkoff et al., 1994) this notion has been made explicit as the fundamental first-tier principle of reference. If children avoid mapping a second label to the visible object, they are left with having heard a novel label being used to refer to something but without an appropriate referent in sight. This, in turn, would lead children to search for a referent for the label.

In other words, by blocking the immediately obvious choice, mutual exclusivity has left the target of a referential label unfilled and should therefore contribute, in conjunction with something like Golinkoff et al.’s principle of reference, to children’s active seeking out of appropriate referents. The lexical gap principle is triggered by the presence of a novel object that creates a gap. Without the object to create the gap, lexical gap filling does not guide children’s behavior. On the lexical gap hypothesis, children require a nameless object to initiate their mapping of the word. On the mutual exclusivity account, children do not need an external incentive to map the novel word—they are motivated by the lexical constraints, in particular mutual exclusivity and the principle of reference, to find an appropriate referent.

In the same way, mutual exclusivity appears different from the N3C principle. According to the N3C principle, children map novel labels onto objects for which they do not yet have a name. The presence of the nameless object is required for the novel name–nameless category principle to operate. Situations where children hear a novel label in the presence of nothing but previously named objects are explicitly outside the scope of N3C (Golinkoff et al., 1992, 1994; Mervis et al., 1994a). N3C furthermore explicitly allows children to make the initial assumption that objects can have
more than one label (Golinkoff et al., 1992). The modification of the indirect word learning task in the present studies thus serves to clarify the distinction between mutual exclusivity and the N3C principle.

The third alternative explanation for previous findings from indirect word learning tasks is that children may be able to determine a speaker’s referential intent through their knowledge of social practices involved in communication and through pragmatic cues (e.g., Bloom, 1998, 2000; Clark, 1997; Diesendruck & Markson, 2001; Tomasello, 2001). In the case of indirect word learning, children base their reasoning on assumptions about speakers’ intent to refer rather than assumptions about how words refer. In part, children may reason along the lines of, “If the speaker meant to refer to the ball, she would have said ‘ball.’” This social-pragmatic account would predict the same pattern of results as mutual exclusivity. Although the present studies were not designed to distinguish between this explanation and mutual exclusivity, we will have more to say about this issue in the discussion.

In sum, the present studies were designed to assess the role of mutual exclusivity in fostering indirect word learning. In particular we examined word learning that takes place without explicit referential cues from the speaker and without a novel object to serve as a ready referent for the new word. The first goal of this work was to determine whether this positive function of mutual exclusivity is available to babies before the time of the naming explosion or whether, as others have argued, it is a later development. The second goal of this work was to determine whether the use of mutual exclusivity can be differentiated from a propensity to fill lexical gaps. Specifically this work addressed whether mutual exclusivity leads 15–20-month-olds to reject second labels for objects and whether it thereby helps motivate them to search for another referent.

2. Study 1

The purpose of Study 1 was to determine whether babies between 15 and 19 months of age are capable of using mutual exclusivity unaided by a bias to fill lexical gaps. If babies attempt to preserve the mutual exclusivity of object labels, then that should motivate them to reject a second label for a known object and help lead them to search for another possible referent of the novel term.

The procedure was to first briefly familiarize babies with a bucket as a location in which to search for unseen objects. Then on each trial a baby viewed a bucket and a familiar object whose label was known to that baby. On Familiar Label trials babies were asked to find the familiar object. On Novel Label trials babies were asked to find the “press” or some other novel object. On baseline No Label trials, children were asked to find “one.” Mutual exclusivity should lead children to indicate the familiar object less and to search more in the Novel Label than in the Familiar Label condition. Children could be searching less when hearing a familiar label because their understanding of the familiar label causes them to inhibit their baseline level of searching. Thus it is important to compare the rate at which children select the familiar object and search for another object when they hear a novel label relative
to their baserate of searching without hearing a label. In sum, the predictions for object selection are that hearing a novel label should motivate children to reject second labels for the familiar object and thus indicate the familiar object less in the Novel Label than in the Familiar Label and No Label conditions. Similarly, the predictions for search are that babies should search more in the Novel Label condition than in the Familiar Label and No Label conditions.

2.1. Methods

2.1.1. Participants

Thirty-two babies participated in this study. Their parents had responded to ads in Palo Alto and San Francisco Bay Area community newspapers and parent-directed publications inviting them to participate in the study. Parents were offered a small gift to thank them for their participation. All of the babies were full-term healthy infants from largely middle-class monolingual English-speaking families. The babies were divided into younger and older age groups designed to span 18 months—the average age of the vocabulary spurt. Sixteen babies were in the younger age group, 8 males and 8 females. They ranged in age from 15 months, 10 days to 17 months, 18 days, with a mean age of 16 months, 6 days. Sixteen babies were in the older age group, 8 males and 8 females. They ranged in age from 18 months, 2 days to 19 months, 28 days, with a mean age of 18 months, 28 days. Two additional babies were run but dropped from the study, one because of fussiness, and one because of equipment failure.

2.1.2. Materials

The materials consisted of an opaque plastic bucket, five novel objects and six familiar objects, plus a ball and a shoe which were used in the familiarization phase. The six familiar objects for each baby were selected from a possible set of 10 including: baby, banana, bottle, car, cat, cup, dog, keys, phone, and spoon. Which familiar toys were selected for a given child depended on which words the child’s parents were most confident that the child understood. As parents were contacted to be scheduled, the experimenter asked them to judge which of the 10 object labels their baby knew and to rate their confidence in this judgment as “sure,” “probably,” “maybe,” “probably not,” and “no.” From this list, six objects were selected to be included in the experimental phase of the study. When possible, only object labels that the parent was sure his or her baby knew were selected. When this was not possible, object labels that the parent judged the baby probably knew were included to bring the total to 6. Two of the five novel objects were used only in the familiarization phase: a flat disk surrounded with suction cups and a half-sphere-shaped noisemaker. The three novel experimental items were a wonton maker (called a “crimp”), a garlic press (called a “press”), and a corn butterer (called a “gadget”).

2.1.3. Procedure

The baby was brought to the experimental room by his or her parent. While the experimenter explained the procedure to the parent, the baby was free to crawl
around the room and play. Once the experimenter and parent judged that the baby was comfortable, the baby was placed in a sassy seat at a small table. The parent sat next to the baby and the experimenter sat across from the baby. Parents were asked to refrain from initiating any interactions with their babies and to avoid pointing at or discussing any of the toys.

Familiarization trials. To familiarize the baby with the bucket as a possible location for objects, each baby was asked to find something in the bucket four times. None of the toys used in this familiarization phase were used in the experiment proper. For two of the trials the experimenter visibly placed a familiar toy in the bucket, say a ball, shook the bucket, and asked the child “Can you find the ball? Where is the ball?” If the child reached for the bucket, the experimenter emptied the toy out of the bucket, allowing the child to play with it for a few seconds. If the child did not reach for the bucket, the experimenter shook it again and repeated the question. If the child still did not reach for the bucket, the experimenter placed the bucket in front of the child and emptied the object out onto the table and allowed the baby to play with it. For two of the trials, this procedure was repeated but with an object for which the baby knew no label. For example, the experimenter visibly placed a flat disk ringed with suction cups in the bucket, shook the bucket and asked the child “Can you find the gripper? Where is the gripper?” Whether the first object in the bucket was familiar or novel was counterbalanced across children. On the following trials, the order of familiar and novel objects alternated. On two trials the bucket was placed on the left and on two trials on the right side of the table. Thus, these four trials provided equal opportunity of seeing familiar and novel objects in the bucket.

Experimental trials. After the brief familiarization, the experimental trials began. On each of the nine experimental trials, the experimenter placed a visible familiar object and a bucket on the table in one of two fixed positions out of reach of the baby—77 cm from the baby, 47 cm apart (aligned with the arms of the sassy seat). The left/right position of the bucket was randomly determined with the constraint that the bucket could be in the same position for not more than five trials out of the nine, and for not more than three consecutive trials. Out of sight of the child, an object had been placed in the bucket. When the bucket was brought into view, the experimenter shook it to let the child know something was there. The visible object and the bucket were placed on the table in a stereotyped way: both of them marching from the edge of the table to the fixed positions. This was done to draw the child’s attention to both the toy and the bucket equally. Once the objects were placed on the table, the experimenter called the child’s name and tapped the table to get the child’s attention. After the experimenter established eye-contact with the child, she asked one of the three experimental questions. During this time and until the baby made a selection, the experimenter avoided looking at the toy or bucket and instead looked at the baby.

In the Familiar Label condition, the experimenter asked the child to find the visible toy (which was always familiar) saying, for example, “Can you find the bottle? Where is the bottle?” The familiar objects were called by the label the parent had reported was most commonly used with the baby. So the doll, for example, might be called “doll,” “dolly,” “baby doll,” or “baby,” depending on what a given child’s
parent would have said. The experimenter requested the object until the baby clearly indicated the visible familiar toy or the bucket, or showed some other obvious signs of searching such as looking on the floor, or was non-responsive after five repetitions. The experimenter then gave the baby both toys to play with. That is, the visible familiar toy was pushed within reach of the baby and the bucket was emptied in front of the child. The baby was allowed to play with the toys for about 10 or 15 s, then the child was asked to hand the toys over to a puppet. The next trial then began.

For the baseline No Label condition, the experimenter said “Can you find one? Where is one?” Again once the baby indicated either the bucket or the visible familiar toy, or seemed to look around for something else, the visible familiar toy was brought within reach of the baby and the bucket emptied in front of the baby. There were three familiar trials and three baseline no label trials. For three of these six trials the bucket toy was a familiar toy for which the baby knew a label, and for three it was an object without a known label.

In the Novel Label condition, the experimenter used a novel term in asking the child to find an object, for example, “Can you find the crimp? Where is the crimp?” As in the other two conditions, once the baby indicated the visible familiar toy or the bucket, the visible familiar toy and the bucket toy were placed in front of the baby to play with. For these novel label trials, the bucket contained a different novel object for each trial.

There were three trials of each of the three conditions. These nine questions were divided into three blocks of three such that each block contained one question from each of the three conditions. Within a block of three, the order of the questions was determined randomly for each participant. The sessions were videotaped with the baby and the objects in full view.

At the conclusion of the session the parent was instructed on how to complete the MacArthur Communicative Development Inventory (Fenson et al., 1991), a language inventory which assesses early vocabulary acquisition. Parents were requested to mail us the completed inventory.

2.1.4. Coding

The videotapes were coded with the coder blind as to which questions babies had been asked. This was accomplished by constructing coding tapes that: (1) had all of the audio information removed, (2) had the segment of each trial that began with eye contact between the experimenter and the baby and ended just before the experimenter revealed the bucket toy, and (3) each coding tape had only one trial from a given baby so that no comparative information across trials was available to infer what condition was being coded for a given baby.

The coder judged whether the baby indicated the visible toy (by reaching or pointing to it), whether the baby indicated the bucket toy (by reaching or pointing towards the bucket), and whether the baby searched the environment (by looking on the floor, glancing from side to side, scanning the room, etc.).

To assess reliability of the coding categories, a second coder independently coded one of the coding tapes which comprised 1/9 of the data. The coding was highly
reliable with agreement across categories ranging from 94 to 97% and averaging 95% overall.

2.2. Results

2.2.1. Vocabulary measures

Our goal in this study was to study children just around the age of the naming explosion. We expected many of the 15–17-month-olds to be just on the verge of the naming explosion while many of the 18–19-month-olds should be in the midst of it. The data from the MacArthur vocabulary questionnaires confirmed that the two age groups represented pre- and post-spurt populations. Questionnaires were returned for 13 children at each age group. The 15–17-month-olds’ reported productive vocabulary ranged from 5 to 160 words with a mean of 46. The 18–19-month-olds’ productive vocabulary ranged from 23 to 297 with a mean of 127. The older babies had significantly larger productive vocabularies than the younger ones, $F(1, 24) = 7.35, p < .05$. The mean productive vocabulary scores for the older babies was significantly greater than 50, a conventional cut-off point for the vocabulary explosion, $t(12) = 2.93, p < .05$, but the younger babies’ mean vocabulary was not significantly less than 50, $t(12) < 1$. On the other hand, the distribution of vocabulary scores established that many of the younger babies had not quite reached the naming explosion, while many of the older ones had. Nine of the 13 older babies had vocabularies greater than 50, while only four of the younger babies did, $\chi^2(1, N = 26) = 3.85, p < .05$.

The first set of analyses was conducted on the number of times that babies indicated the visible familiar toy in the three conditions. Table 1 presents these data converted to mean percent of trials on which babies selected the visible familiar toy. The data were analyzed by an Age (Old vs. Young) $\times$ Condition (Familiar Label, Novel Label, No Label) ANOVA, with condition as a within-subjects variable. There was a highly significant effect of condition, $F(2, 60) = 18.56, p < .001$. There was no main effect of age, $F(1, 30) = .2$, but there was an Age $\times$ Condition interaction, $F(2, 60) = 3.52, p < .05$. This interaction is that the older babies picked the visible familiar toy slightly less than the younger babies in the Novel Label condition and in the No Label baseline condition but more than the younger babies in the Familiar Label condition. Planned comparisons were used to test whether babies showed

<table>
<thead>
<tr>
<th>Age (months)</th>
<th>Condition</th>
<th>Familiar Label</th>
<th>Novel Label</th>
<th>No Label</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
</tr>
<tr>
<td>16</td>
<td></td>
<td>73</td>
<td>33</td>
<td>44</td>
</tr>
<tr>
<td>19</td>
<td></td>
<td>90</td>
<td>20</td>
<td>35</td>
</tr>
</tbody>
</table>
evidence of using mutual exclusivity in this situation. The analyses were conducted separately for each age group.

One clear prediction from the hypothesis that very young children use mutual exclusivity to narrow their interpretations of novel labels is that children should be less inclined to indicate a familiar toy as a referent of a novel label compared to the known label. This was borne out for both age groups. The older babies indicated the familiar toy 2.69 times out of a possible 3 (90%) when asked for the familiar toy and only 1.06 times out of 3 (35%) when they heard a novel label, paired- \( t(15) = 7.34, p < .001 \). Similarly, the younger babies indicated the familiar toy 2.19 times out of 3 (73%) when asked to find familiar toys and only 1.31 times (44%) when they heard the novel labels, paired- \( t(15) = 2.41, p < .05 \). A sign test confirmed these findings. Twenty-five babies picked the visible familiar toy more in the Familiar Label than in the Novel Label condition compared to only two babies who did the reverse, \( p < .005 \).

A second prediction from the hypothesis that very young children use mutual exclusivity is that upon hearing a novel label children should select a familiar object as its potential referent less often than their baseline level of selecting familiar objects. This was again true of both age groups, though more clearly for the younger babies. The older babies indicated the familiar toy 1.06 times (35%) in the Novel Label condition and 1.63 times (54%) in the baseline condition, \( t(15) = 1.95, p = .07 \). The younger babies selected the familiar toy 1.31 times out of 3 (44%) in the Novel Label condition and 2.19 times (73%) in the No Label baseline condition, paired- \( t(15) = 3.05, p < .01 \). These differences were significant by a sign test as well. Overall 14 babies selected the visible familiar toy more in the No Label baseline condition than in the Novel Label condition and only two babies selected the reverse, \( p < .005 \).

In sum, both predictions were supported: while looking at a bottle, for example, and asked to find, e.g., a “crimp,” babies pick the bottle less often than when asked for a bottle and less often than their baseline level when asked to find “one.” Thus, as predicted, upon hearing a novel label, 15–19-month-old babies are led to reject a familiar toy as a possible referent of the novel term.

Mutual exclusivity helps babies avoid mapping errors and it can also help lead them to search for a novel object as a potential referent for the novel label. To determine if this function of mutual exclusivity is available to babies we examined the number of times babies either indicated the bucket or searched in some other way, such as scanning the room or leaning over to look on the floor. Although this analysis is not independent of the first, it is not entirely redundant. That is, babies might fail to pick the familiar toy but also fail to search for another object. They could just be confused or non-responsive on a given trial and not be credited with either response. Conversely babies might sometimes simultaneously indicate both the visible toy and the bucket.

The mean percent of times babies appeared to search for an object other than the visible familiar object is presented in Table 2. The number of times babies appeared to search for an object was analyzed by an Age (Older vs. Younger) \( \times \) Condition (Familiar Label, Novel Label, No Label) ANOVA, with condition being a within-subjects variable. There was a significant main effect of condition, \( F(2, 60) = \ldots \)
18.75, \( p < .001 \). There was no main effect of age, \( F(1, 30) = .35 \), nor did age interact with condition, \( F(2, 60) = 2.18, \ p = .12 \). Even though there was no interaction with age, we analyzed the results separately for each age group to double check whether the findings were reliable at each age.

The first prediction from the mutual exclusivity hypothesis is that babies should search for an alternative object as a possible referent more in the Novel Label than in the Familiar Label condition. This prediction was supported for both the 18–19-month-olds and the 15–17-month-olds. The older babies showed signs of searching 1.81 times out of 3 (60%) in the Novel Label condition and only .31 times (10%) in the Familiar Label condition, paired-\( t(15) = 6.71, \ p < .001 \). The younger babies showed signs of search 1.56 times (52%) in the Novel Label condition compared to .62 times (21%) in the Familiar Label condition, paired-\( t(15) = 2.91, \ p < .05 \). A sign test confirmed these findings with 27 babies searching more in the Novel Label than in the Familiar Label condition with only one baby showing the reverse pattern, \( p < .002 \).

The second prediction is that babies should show more signs of searching for a novel object in the Novel Label condition than in the No Label baseline condition. This prediction was supported for the younger babies but did not reach significance for the older. The weaker finding for the older babies is caused not by any drop in their searching in the Novel Label condition but by their higher baserate preference for looking in the bucket even in the baseline No Label condition. The 18–19-month-olds searched a mean of 1.81 times out of 3 (60%) in the Novel Label condition compared to 1.31 (44%) in the No Label condition, paired-\( t(15) = 1.65, \ p = .12 \). The 15–17-month-olds searched 1.56 times out of a possible 3 (52%) in the Novel Label condition compared to .81 times (27%) in the No Label condition, paired-\( t(15) = 2.82, \ p < .05 \). Non-parametric analyses confirmed the prediction: Overall 12 babies searched more in the Novel Label than in the No Label condition, with four showing the reverse pattern, \( p < .05 \), by a sign test.

Although data on vocabulary size was not reported for all the babies, the available data suggest that the effects hold up both for babies who have not yet entered the vocabulary spurt and for those who are in the midst of it. Table 3 presents the percentage of trials on which babies selected the visible toy with a known label in each condition. The babies are divided into two groups: Those with productive vocabularies of less than 50 words and those with vocabularies of 50 or more words, regardless of age. These data were analyzed by a Vocabulary Level (Less than 50, 50 or more) \( \times \) Condition (Familiar, Novel, and No Label) ANOVA. As expected
condition was highly significant, $F(2, 48) = 14.50$, $p < .0001$. Moreover, there was no main effect for vocabulary level, $F(1, 24) = .11$, nor did vocabulary level interact with condition, $F(2, 48) = 1.44$, $p = .27$. Thus the overall effect of condition can be generalized to both pre- and post-spurt babies. Individual t tests were used to examine if the two predictions from mutual exclusivity held up for both pre- and post-spurt babies. The first prediction compares babies’ selection of a familiar toy as a referent of a novel label compared to the known label for a toy. Pre-spurt babies indicated the familiar toy 2.42 times of a possible 3 (81%) when asked for the familiar toy and 1.42 times out of 3 (47%) when they heard a novel label, paired-$t(11) = 3.63$, $p < .005$. Similarly, post-spurt babies indicated the familiar toy 2.50 times (83%) in the Familiar Label condition and .85 times (29%) in the Novel Label condition, paired-$t(13) = 3.97$, $p < .005$. Thus, this prediction was confirmed for both vocabulary groups for these selection data. The second prediction compares babies’ selection of a familiar toy when they hear a novel label with a baseline condition of selecting a familiar toy when no label is heard. The pre-spurt babies selected the familiar toy 1.42 times (47%) when they heard a novel label and 1.75 times (58%) when no label was heard, paired-$t(13) = 1.17$, $p = .26$. The post-spurt babies selected the familiar toy .86 times (29%) when they heard a novel label and two times (67%) out of three in the No Label condition, paired-$t(13) = 3.17$, $p < .01$. Thus the predicted difference against baseline was confirmed for the post-spurt babies but not the pre-spurt babies.

An analogous analysis was carried out on the search measure. Table 4 presents the percentage of trials on which pre- and post-spurt babies searched, either in the bucket or some other location, for each condition. These data were analyzed by a Vocabulary (Less than 50, 50 or more) × Condition (Familiar, Novel, and No Label) ANOVA. Again condition was highly significant, $F(2, 48) = 13.75$, $p < .0001$.

Table 3
Mean percent of trials in Study 1 on which babies selected a familiar toy as a referent of a novel label compared to the known label for a toy.

<table>
<thead>
<tr>
<th>Vocabulary level</th>
<th>Condition</th>
<th>Familiar Label</th>
<th>Novel Label</th>
<th>No Label</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>Less than 50 words ($N = 12$)</td>
<td>81</td>
<td>22</td>
<td>47</td>
<td>33</td>
</tr>
<tr>
<td>50 or more words ($N = 14$)</td>
<td>83</td>
<td>31</td>
<td>29</td>
<td>32</td>
</tr>
</tbody>
</table>

Table 4
Mean percent of trials in Study 1 on which babies searched according to vocabulary level.

<table>
<thead>
<tr>
<th>Vocabulary level</th>
<th>Condition</th>
<th>Familiar Label</th>
<th>Novel Label</th>
<th>No Label</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>Less than 50 words ($N = 12$)</td>
<td>14</td>
<td>22</td>
<td>50</td>
<td>30</td>
</tr>
<tr>
<td>50 or more words ($N = 14$)</td>
<td>17</td>
<td>31</td>
<td>64</td>
<td>33</td>
</tr>
</tbody>
</table>
And again there was no main effect of vocabulary level, $F(1, 24) = .07$ and no interaction of vocabulary level with condition, $F(2, 48) = 1.20, p = .31$. The pattern of the individual comparisons by paired $t$ tests paralleled the significance levels reported above for the data from the selection of the visible toy.

Whether the babies in the study were grouped by age or vocabulary, the predicted difference between the Familiar Label and the Novel Label conditions was upheld for all groups. The predicted differences against baseline, however, was confirmed only for younger babies and did not quite reach significance for the older babies. This contrasts with the results when babies were grouped by vocabulary, where the prediction was upheld for the high vocabulary group but did not reach significance for the low vocabulary group. At first sight, these findings might suggest that only young and only post-spurt babies use mutual exclusivity to guide search—an odd conclusion given that age and vocabulary status are positively, not negatively, correlated. A closer inspection of the data reveals that the data in the Novel Label condition remains essentially unchanged whether grouped by age or vocabulary whereas the data in the No Label baseline condition changes direction as a function of age. Examining babies’ behavior by age in the No Label baseline condition, Table 2 shows that older babies as a group prefer to search more in the bucket (44%) than younger babies (27%). Consequently the older babies, as shown in Table 1, tend to select the visible toy less (54%) than the younger babies (73%). Recall that the vocabulary data showed that four older children had productive vocabularies of less than 50 words and four younger children had vocabularies of 50 words or more. As the four older children who are inclined to search are moved to the low vocabulary group and the five younger children who are less inclined to search are moved to the high vocabulary group, the baseline levels for searching and selecting the visible toy in the vocabulary data are pulled in the opposite direction of those in the age data. In other words, when asked to “show me one” older babies are more interested in seeing what is in the bucket. We are arguing that this reflects their age and exploratory behavior, not their vocabulary acquisition. The primacy of age over vocabulary level on babies’ search behavior in the No Label baseline condition was confirmed with a regression analysis. The percentage of trials on which babies searched on the baseline trials was entered as the dependent variable and vocabulary and age as independent dichotomous variables. Age significantly predicted searching, $\beta = .46, p < .05$, while vocabulary level did not, $\beta = -.30, p = .14$. This was also the case when babies’ selection of the visible toy was entered as the dependent variable. Age was significantly negatively related to selection, $\beta = -.50, p < .05$, while vocabulary was unrelated, $\beta = .28, p = .16$. Regressions were also performed for the other two dependent variables. Neither age nor vocabulary were significant predictors in the Familiar label or the Novel Label conditions. Thus, while it may appear as if the predicted difference between search in the Novel Label condition relative to baseline is dependent on the size of babies’ vocabulary, the data show that it likely reflects age differences in babies’ search behavior in the baseline condition rather than inherent differences in their use of mutual exclusivity according to vocabulary size. This interpretation will be given further support by a related finding in Study 3.
To summarize, this study was designed to determine whether babies around the average age of the naming explosion are capable of using mutual exclusivity to guide their interpretation of novel labels. According to mutual exclusivity babies should reject second labels for objects. In order to avoid having children use an alternative strategy of filling lexical gaps, they were prevented from seeing any novel objects at the time the label was heard. Babies in this study learned that a bucket was a possible location in which to find objects. They then viewed a familiar object and an opaque bucket whose contents were unknown. To take one concrete example, a baby might see a toy dog and the bucket placed on the table. They were then asked to find an object in one of three ways, either using the familiar label, “Can you find the dog?” a novel label, “Can you find the crimp?” or no label “Can you find one?” As predicted under the assumption that babies are rejecting second labels for objects, babies indicated the visible familiar object less when hearing a novel label than when hearing the object’s familiar label, and less than their baserate of picking the familiar object. Conversely they searched for something else as a potential referent more when hearing the novel label than when hearing the familiar label and at least the younger babies more than their baseline level of searching. Thus babies as young as 15 months old were found to honor the mutual exclusivity assumption, unaided by a bias to fill lexical gaps.

Study 2 was designed to probe the limits of when mutual exclusivity will lead babies to reject second labels for objects and search for other possible referents. In Study 1, a familiar object and a bucket as a potential location for objects were both placed out of reach of children as the experimenter asked them to find an object. In Study 2, babies were given a familiar object to play with while the bucket remained out of reach. This was done to make the familiar object more compelling and more salient to the babies at the time the novel label was heard. The question was whether mutual exclusivity is powerful enough (1) to lead children to reject a second label for an object they are playing with and (2) to interrupt their play to search for a novel object as a potential referent of the novel label.

3. Study 2

3.1. Methods

3.1.1. Participants

Thirty-two babies participated in this study. Their parents had responded to ads in Palo Alto and San Francisco Bay Area newspapers inviting them to participate in the study. Parents were given a small gift for their participation. All of the babies were full term healthy infants from monolingual English-speaking families. The babies were divided into younger and older age groups designed to span 18 months—the average age of the vocabulary spurt. Sixteen babies were in the younger age group, 8 males and 8 females. They ranged in age from 15 months, 5 days to 17 months, 25 days, with a mean age of 16 months, 3 days. Sixteen babies were in the older age group, 8 males and 8 females. They ranged in age from 18 months,
1 day to 20 months, 15 days, with a mean age of 19 months, 5 days. Three additional babies were run but dropped from the study, two because of fussiness, and one because of equipment failure.

3.1.2. Materials

The materials were identical to that of Study 1, including the familiar and novel objects and an opaque bucket.

3.1.3. Procedure

The procedure was identical to that of Study 1 with one exception. In Study 1, both the visible familiar toy and the opaque bucket were out of reach of the baby at the time the experimenter asked the baby to, e.g., “find the dog,” “find the crimp,” or “find one.” In Study 2, the visible familiar object was given to the child to play with but the bucket was kept out of reach. So when the child was asked to find an object, they were already engaged in play with the visible familiar toy.

3.1.4. Coding

As in Study 1, coders were blind as to what question a baby had been asked. This was accomplished by constructing coding tapes according to the procedure used in Study 1.

Coders judged whether the baby indicated the visible toy, the bucket toy, or otherwise searched the environment (such as looking around the room, on the floor, etc.). As in Study 1, babies indicated the bucket toy by reaching for it or pointing to it. But because the babies were playing with the visible toy, they could also hold it up or touch it as well as pointing to it. Because the possible responses were somewhat more varied, we had the coder rate her confidence in her decision.

Again, the coding categories were highly reliable. One indication that the coding was straightforward was the primary coder’s confidence in the judgments made. On a scale from 1 to 5 where 5 is most confident, the average confidence across categories was 4.7. In particular, average confidence was high (4.6) for judgments about whether or not babies indicated the toy they were playing with.

A second coder independently coded one of the coding tapes or 1/9 of the data. Agreement across categories ranged from 94 to 100%, and averaged 97%.

3.2. Results

3.2.1. Vocabulary measures

We received MacArthur vocabulary questionnaires for 12 of the younger babies and 13 of the older ones. The range in reported productive vocabulary was 6–196 for the 15–17-month-olds, with a mean of 67 words. The range in reported productive vocabulary was 18–20 to 406 for the 19-month-olds with a mean of 162 words. Although the mean vocabulary scores were different for the two age groups, $t(17) = 2.19, p < .05$, the distributions were not different when 50 words was taken as the cut-off point. Six of the 12 15–17-month-olds had vocabularies of less than 50 words, while four of the 13 18–20-month-olds did.
The first set of analyses was conducted on the number of times that babies indicated the visible familiar toy in the three conditions. Table 5 presents these data. The data were analyzed by an Age (Old vs. Young) × Condition (Familiar Label, Novel Label, No Label) ANOVA, with condition as a within-subjects variable. There was a highly significant effect of condition, $F(2, 60) = 13.12$, $p < .001$. There was no main effect of age, $F(1, 30) = 1.82$, $p = .19$, nor did age interact with condition, $F(2, 60) = 1.31$, $p = .26$. Although there was no interaction with age, the results were analyzed separately for each age group to determine whether the effects were reliable at each age.

The first analyses tested the prediction from mutual exclusivity that children should indicate the familiar toy they were playing with less when asked to find something with a novel label than when the familiar label was used. This was confirmed for both age groups. The older babies indicated the familiar object 2.75 times (92%) in the Familiar Label condition compared to 1.56 (52%) times in the Novel Label condition, paired-$t(15) = 3.88$, $p < .005$. The younger babies indicated the familiar object 2.69 (90%) times in the Familiar Label condition compared to 2.06 (69%) times in the Novel Label condition, paired-$t(15) = 3.10$, $p < .01$. Overall, 23 of the babies selected the familiar toy more in the Familiar Label condition than the Novel Label condition, while four showed the reverse pattern, $p < .005$, by a sign test.

A second set of analyses examined the prediction that babies should indicate the familiar toy as a referent less in the Novel Label condition than in the No Label baseline condition. This prediction was confirmed for both age groups combined. The results were not significant for both ages taken separately, however. The older babies indicated the familiar toy they had in hand 1.56 (52%) times in the Novel Label condition compared to 2.25 (75%) times in the No Label condition, paired-$t(15) = 2.03$, $p = .06$. The younger babies indicated the familiar toy 2.06 (69%) times in the Novel Label condition and 2.38 (79%) times in the No Label condition, paired-$t(15) = 1.43$, $p = .17$. The prediction was confirmed with both age groups combined: babies picked the familiar toy 1.81 (60%) times in the Novel Label condition versus 2.31 (77%) times in the No Label condition, paired-$t(31) = 2.49$, $p < .05$, and as reported earlier there was no interaction with age. Non-parametric analyses confirmed this finding: Overall, 16 babies picked the familiar toy less in the No Label condition than the Novel Label condition, while six showed the reverse pattern, $p = .052$ by a sign test.

Table 5
Mean percent of trials in Study 2 on which the visible familiar toy was indicated according to age

<table>
<thead>
<tr>
<th>Age (months)</th>
<th>Condition</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Familiar Label</td>
<td>Novel Label</td>
<td>No Label</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
</tr>
<tr>
<td>16</td>
<td>90</td>
<td>16</td>
<td>69</td>
<td>23</td>
<td>79</td>
</tr>
<tr>
<td>19</td>
<td>92</td>
<td>15</td>
<td>52</td>
<td>32</td>
<td>75</td>
</tr>
</tbody>
</table>
Thus the mutual exclusivity assumption helped 15–20-month-old babies reject a familiar object as a referent of the novel term even though they were holding or playing with the object at the time the novel label was heard. They indicated the familiar object less in the Novel Label condition than in either the Familiar Label or the baseline Novel Label condition. Under these more taxing circumstances, the results were not always as striking as those of Study 1.

Analyses were also conducted on the number of times babies showed signs of searching for another object, either by reaching to see what was in the bucket or by looking around the room, etc. As in Study 1, the search data are not independent of the data on reaching for the visible toy, but they are not entirely redundant as on a given trial, a child could exhibit both behaviors or neither behavior. Table 6 presents the percent of time babies searched in each condition.

An Age (Old vs. Young) × Condition (Familiar Label, Novel Label, No Label) ANOVA was used to analyze the search data, with condition as a within-subjects variable. There was a significant effect of condition, $F(2, 60) = 8.58, p < .001$. Although there was no significant effect of age, $F(1, 30) = 2.35, p = .14$, age did interact with condition, $F(2, 60) = 10.99, p < .001$. As will be verified in the analyses to be reported next, older babies searched more in the Novel Label condition than in the other two conditions as predicted, but the younger babies did not.

The first set of planned comparisons examined the prediction from mutual exclusivity, that babies should search more in the Novel Label condition than in the Familiar Label condition. Although this prediction was very clearly supported for the older babies, it was not for the younger babies. The 18–20-month-olds searched a mean of .44 (15%) times in the Familiar Label condition but a full 2.06 (69%) times in the Novel Label condition, paired-$t(15) = 7.34, p < .001$. The younger babies showed very little explicit search in this study, searching a mean of .81 (27%) times in the Familiar Label condition and .75 (25%) times in the Novel Label condition, paired-$t(15) < 1$. Apparently the 15–17-month-olds were not willing to disrupt their play with the toy they had in hand to search for something else. Non-parametric analyses confirmed this pattern for the older babies: Fifteen of the older babies searched more in the Novel Label condition than in the Familiar Label condition, with no babies showing the reverse pattern, $p < .005$ by a sign test.

The second set of analyses examined the prediction that babies should search more in the Novel Label condition than in the No Label condition. Here again, the prediction was clearly supported for the older but not the younger babies. The younger babies searched .75 (25%) times in the Novel Label condition and 1.06

---

Table 6
Mean percent of trials in Study 2 on which babies searched according to age

<table>
<thead>
<tr>
<th>Age (months)</th>
<th>Condition</th>
<th>Mean</th>
<th>SD</th>
<th>Mean</th>
<th>SD</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Familiar Label</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>27</td>
<td>28</td>
<td></td>
<td>25</td>
<td>31</td>
<td>35</td>
<td>31</td>
</tr>
<tr>
<td>19</td>
<td>15</td>
<td>17</td>
<td></td>
<td>69</td>
<td>28</td>
<td>35</td>
<td>31</td>
</tr>
</tbody>
</table>
(35%) times in the No Label baseline condition, paired-$t(15) = 1.04$, $p = .32$. Again the babies may have been unwilling to stop playing with the toy in hand long enough to search. In contrast, the older babies searched 2.06 (69%) times in the Novel Label condition compared to 1.00 (33%) times in the No Label condition, paired-$t(15) = 3.78$, $p < .005$. Eleven of the older babies searched more in the Novel Label than in the No Label condition with only one baby showing the reverse pattern, $p < .05$, by a sign test.

Another way of analyzing the data is to look at children’s performance as a function of their vocabulary level instead of age. As in Study 1, babies were divided into two groups—those with productive vocabularies of less than 50 words ($n = 11$) and those with 50 or more words ($n = 14$). The percent of times babies selected the visible toy is presented in Table 7 as a function of vocabulary level and condition. The data were analyzed by a Vocabulary (Less than 50, 50 or more) $\times$ Condition (Familiar, Novel, and No Label) ANOVA, with condition a within-subjects variable. There was a strong effect of condition, $F(2, 46) = 11.14$, $p < .001$, but no effect of vocabulary level, $F(1, 23) = .03$, nor any interaction of vocabulary with condition, $F(2, 14) = .30$. Thus the overall condition differences found generalized across both high and low vocabulary levels.

The search data were also analyzed dividing the babies by vocabulary level instead of age. Table 8 shows how often (converted to percentages) babies searched as a function of their vocabulary level and condition. A Vocabulary (Less than 50, 50 or more) $\times$ Condition (Familiar, Novel, and No Label) ANOVA with condition as a within-subjects variable was used to analyze the mean search scores. There was a significant effect of condition, $F(2, 46) = 5.59$, $p < .01$, but no effect of vocabulary level, $F(1, 23) = .002$, nor any interaction of vocabulary and condition.

Table 7
Mean percent of trials in Study 2 on which babies indicated the visible familiar toy according to vocabulary level

<table>
<thead>
<tr>
<th>Vocabulary level</th>
<th>Condition</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Familiar Label</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>77</td>
<td>17</td>
<td>61</td>
<td>20</td>
<td>77</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Novel Label</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>93</td>
<td>14</td>
<td>55</td>
<td>33</td>
<td>74</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No Label</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 8
Mean percent of trials in Study 2 on which babies searched according to vocabulary level

<table>
<thead>
<tr>
<th>Vocabulary level</th>
<th>Condition</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Familiar Label</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>27</td>
<td>20</td>
<td>49</td>
<td>38</td>
<td>39</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Novel Label</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>19</td>
<td>17</td>
<td>52</td>
<td>36</td>
<td>43</td>
<td>38</td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No Label</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Thus the overall effect of condition generalized across babies of both high and low vocabularies. We do not present further analyses of the vocabulary data past the omnibus ANOVA. Evidence from Study 1 showed that grouping the data by vocabulary level is influenced by babies’ tendency to search in the bucket at different age levels. Babies’ behavior in Study 2 is further influenced by their willingness to ignore the visible object in their hands. Since age and vocabulary are imperfectly correlated, these factors make it difficult to discern the influence of age vs. vocabulary when the data is grouped by vocabulary. The number of participants in the study does not allow dividing a given age group into subgroups of high and low vocabulary without loss of statistical power.

The failure of the younger babies to search in this study may be a function of the task demands of inhibition as well as their lexical principles. In Study 1 with both the visible toy and the bucket toy out of reach, the younger babies requested the bucket toy as predicted on the mutual exclusivity hypothesis. Hearing a novel label caused these 15–17-month-olds to reject a familiar toy as an appropriate referent and search for another potentially more appropriate referent for the new word. When babies were given the visible toy to play with in Study 2, only the 18–20-month-olds searched as predicted upon hearing a novel label relative to both a familiar label and no label. The 15–17-month-olds did not search for another referent in either of those cases.

To summarize, both younger and older babies tended to perform as predicted when their acceptance of a label was measured by how often they indicated the visible toy upon hearing a novel label relative to a familiar label. The predicted difference relative to the baseline No Label condition was obtained only for the older children. Thus even in this more difficult situation where the salience of the toy with a known label was heightened by allowing children to play with it, children used mutual exclusivity to guide their interpretation of a novel term to some extent. But in this situation only the 18–20-month-olds searched for another referent. One likely explanation for the discrepancy between the selection and search measures for the 15–17-month-olds is that to search for another potential referent would have required a more sustained interruption or inhibition of their play and, although the older babies accomplished this, the younger babies were not sufficiently motivated or able to do so.

4. Study 3

Study 3 further explored the limits of mutual exclusivity to motivate search for a referent by removing the obvious place for babies to search. Studies 1 and 2 have established that 15–20-month-olds can use mutual exclusivity to reject second labels for objects and, at least under some circumstances, to search for an appropriate referent for a novel word. In both of these studies a bucket was made available as a potential container for another object. In naturally occurring situations sometimes there will be such an obvious location to search, but sometimes not. In this next study, we explored whether babies around the age of the naming explosion would
be motivated by mutual exclusivity to reject a second label for an object and search for a referent for a novel term even when there was no particularly obvious place to search. The procedure was similar to that of Studies 1 and 2 except that no bucket was used and no novel objects were ever presented to the children. The main question was whether babies would show spontaneous search of the room.

4.1. Methods

4.1.1. Participants

The participants in this study were 36 healthy, full-term, and monolingual English-speaking babies whose parents responded to ads in local newspapers and parent-directed publications. Parents were offered a small gift for their participation. The younger age group consisted of 18 babies ranging in age from 15 months 7 days to 15 months 28 days with a mean of 15 months 18 days. The older group consisted of 18 babies ranging in age from 18 months 1 day to 18 months 25 days with a mean of 18 months 13 days.

4.1.2. Materials

The materials consisted of 12 objects that are among the most likely objects to have known names by very young children: baby, dog, cat bottle, duck, car, keys, book, spoon, shoe, and banana. As in Studies 1 and 2, before the start of the study, each parent was asked to rate whether their baby understood each of these words. For each baby, the experimenter selected three objects representing three of the words that the baby was most familiar with. An attempt was made to vary the items as much as possible across babies.

4.1.3. Procedure

Throughout the procedure babies were seated in a sassy seat at the table and their parent was seated off to the side.

The purpose of this study is to determine if babies will search for an unseen object as the referent of a novel term to avoid accepting a second label for an object with a well-known name. Once a baby scans the room, it will become obvious that there are no other candidate objects around as potential referents for a novel term which may influence their scanning on subsequent trials. For this reason, we had to severely limit the number of trials each baby received. As in the earlier studies we had three kinds of questions: Familiar Label, Novel Label, and No Label baseline, but in this study babies heard only one question of each type.

For each of the questions the experimenter first gave the baby a familiar object with a well-known label and allowed the baby to play with it for a brief time. The experimenter then retrieved the toy (sometimes having the baby give it to a puppet) and placed it on the table in a fixed position out of reach of the baby. The experimenter then tapped on the table, called the baby’s name, and once she established eye-contact with the baby she asked one of the three questions.

In the Familiar Label condition, the experimenter used the correct label for the toy, repeating the question several times. If a spoon was placed on the table, the
experimenter would say “Can you show me the spoon? Where’s the spoon? Find the spoon.”

In the Novel Label condition a novel term was used in the request, but otherwise the procedure was identical to that in the Familiar Label condition. For example, a baby might see a shoe placed on the table out of reach and then be asked, “Can you show me the toma? Where’s the toma? Find the toma.”

In the No Label condition, designed to assess baseline levels of searching, the request was made without using any labels. Otherwise the condition was identical to the other two conditions. A baby might see a banana on the table and be asked “Can you show me one? Where is one? Find one.”

The order of presentation of the Familiar Label, Novel Label, and No Label questions was counterbalanced across children such that each question occurred in the first, second, and third position equally often.

As in Studies 1 and 2, parents were asked to complete the MacArthur Communicative Development Inventory (Fenson et al., 1991) and mail it back to us within the next few days.

4.1.4. Coding

The videotaped responses of babies were coded without the soundtrack so that the coder had no knowledge of what had been requested. Coding began from the time the experimenter requested the object (which was determined in a preliminary coding of the tapes with the soundtrack on). The coder judged whether the baby indicated the visible toy (by reaching or pointing), whether the baby searched the environment (by looking around the room, looking on the floor, looking from side to side, etc.), or whether the baby made some other response such as prolonged looking at the experimenter. After each judgment the coder was asked to rate her confidence on a 5-point scale where 5 is most confident.

The coding system was highly reliable. First, the primary coder was very confident of the judgments made. Overall categories the mean confidence in the judgments was 4.87 on a 5-point scale. For the search category alone the mean confidence was 4.89 indicating that babies made quite identifiable search responses that were readily coded. A second coder independently judged 33% of the data. The coders agreed on 95% of the judgments.

4.2. Results

4.2.1. Vocabulary measures

As expected the 15 and 18-month-olds straddled both sides of the naming explosion. Fifteen of the 18 parents of 15-month-olds completed the MacArthur vocabulary questionnaire. The 15-month-olds productive vocabulary ranged from 0 to 62 words, with a mean of 24.1 which was significantly less than 50, \( t(14) = 5.19, p < .001 \). Fourteen parents of the 18-month-olds completed the questionnaires. The reported productive vocabulary of the 18-month-olds ranged from 20 to 385 with a mean of 121.4 which is significantly greater than 50, \( t(13) = 2.30, p < .05 \). The 15-month-olds had a significantly smaller vocabulary than the 18-month-olds,
The distribution of vocabulary scores less than 50 versus 50 or more differed significantly by age. Thirteen of the 15 15-month-olds had vocabularies of less than 50, while only four of the 14 18-month-olds did, \( \chi^2(1, N = 29) = 10.07, p < .01 \).

In this study, the only object in view at the time babies were asked to retrieve something was a familiar object with a well known name. In particular, there was no bucket to serve as a potential location for a novel object. At one point, the experimenter used a novel label in making the request, e.g., “Can you show me the toma?” The main question the study addressed was whether the preference for keeping category terms mutually exclusive would motivate babies to search their environment for a candidate object to label.

The two predictions from the hypothesis that babies are using mutual exclusivity here is that: (1) they should search the environment for a novel object more when they hear a novel word used in the request than when they hear the familiar label for the visible object and (2) more than when they are simply asked to find “one.” The percentages of children at each age who searched in each of the conditions is presented in Table 9. Because there was only one binary response per condition for each child, the data were analyzed with a categorical repeated measures ANOVA using the SAS CATMOD procedure followed by non-parametric planned comparisons. An Age (Older vs. Younger) × Condition (Familiar Label, Novel Label, and No Label) categorical repeated measures ANOVA with condition as a repeated measure was conducted to examine whether there were differences in the frequency with which babies searched across conditions. There was a significant main effect of condition, \( \chi^2(2, N = 36) = 21.77, p < .0001 \). There was no main effect of age, \( \chi^2(1, N = 36) = .13, p = .72 \), nor did age interact with condition, \( \chi^2(2, N = 36) = 2.62, p = .27 \). We also analyzed the results separately for each of the predictions from mutual exclusivity.

The first prediction from the mutual exclusivity hypothesis is that babies should search for an alternative object as a possible referent more in the Novel Label than in the Familiar Label condition. This prediction was supported for both the 18-month-olds and the 15-month-olds. In the older group, 15 babies (83%) showed signs of searching in the Novel Label condition compared with 9 babies (50%) in Familiar Label condition. The means of each condition differed significantly by a Cochran test, \( Q \) with 1 \( df = 6.00, p < .05 \). In the younger group, 16 babies (89%) showed signs of searching in the Novel Label condition compared to 10 babies (56%) in the Familiar Label condition, \( Q \) with 1 \( df = 6.00, p < .05 \).

### Table 9
Percentages of babies in Study 3 who searched according to age

<table>
<thead>
<tr>
<th>Age (months)</th>
<th>Condition</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Familiar Label</td>
<td>Novel Label</td>
<td>No Label</td>
</tr>
<tr>
<td>15</td>
<td>56</td>
<td>89</td>
<td>50</td>
</tr>
<tr>
<td>18</td>
<td>50</td>
<td>83</td>
<td>72</td>
</tr>
</tbody>
</table>
Babies could be searching more in the Novel Label condition than in the Familiar Label condition because they are rejecting second labels for the objects with well-known labels and searching for another possible referent. Alternatively, they could be searching less in the Familiar condition because the presence of the familiar label causes them to inhibit their baseline level of searching. Thus, the prediction is that babies should search more in the Novel Label condition than their baseline level of searching in the No Label baseline condition. This prediction was clearly supported for the younger babies but not the older babies. Of the 18-month-olds, 15 babies (83%) searched in the Novel Label condition compared to 13 babies (72%) in the No Label condition, $Q$ with $1 \ df = .67$, $p = .41$. Of the 15-month-olds, 16 babies (89%) searched in the Novel Label condition compared to nine babies (50%) in the No Label condition, $Q$ with $1 \ df = 7.00$, $p < .01$.

Vocabulary data were reported for 29 babies and indicated that the predictions were generally held up both for babies who had not yet entered the vocabulary spurt and those who were in the middle of it. The babies were divided into two groups based on their productive vocabulary, with 50 words as the cut-off. Table 10 shows the percentages of children at each vocabulary level who searched in each of the conditions. Whether babies searched was analyzed with a Vocabulary (Less than 50, 50 or more) $\times$ Condition (Familiar Label, Novel Label, and No Label) categorical repeated measures ANOVA with condition as a repeated measure. There was a significant main effect of condition, $\chi^2(2, N = 29) = 18.93$, $p < .0001$, with no main effect of vocabulary, $\chi^2(1, N = 29) = .74$, $p = .39$, and no interaction between vocabulary and condition, $\chi^2(2, N = 29) = .54$, $p = .76$. Planned comparisons showed that the predicted difference for the Familiar and Novel Label conditions was found at both vocabulary levels. For the post-spurt babies, 10 babies (83%) searched in Novel Label condition compared with 5 babies (42%) in the Familiar label condition, $Q$ with $1 \ df = 6.00$, $p < .05$. For the pre-spurt babies, 16 babies (94%) searched in the Novel Label condition compared with 10 babies (59%) in the Familiar Label condition, $Q$ with $1 \ df = 6.00$, $p < .05$. The predicted difference between the Novel Label and baseline No Label conditions was found for the pre-spurt but not the post-spurt babies. For the post-spurt babies, 10 babies (83%) searched in the Novel Label condition compared with 8 babies (67%) in the No Label condition, $Q$ with $1 \ df = .67$, $p = .41$. For the pre-spurt babies, 16 babies (94%) searched in the Novel Label condition compared with 11 babies (65%) in the No Label condition, $Q$ with $1 \ df = 5.00$, $p < .05$. The individual comparisons when the data were divided by vocabulary thus paralleled those when the data were divided by age.

<table>
<thead>
<tr>
<th>Vocabulary level</th>
<th>Condition</th>
<th>Familiar Label</th>
<th>Novel Label</th>
<th>No Label</th>
</tr>
</thead>
<tbody>
<tr>
<td>Less than 50 words ($N = 12$)</td>
<td>59</td>
<td>94</td>
<td>65</td>
<td></td>
</tr>
<tr>
<td>50 or more words ($N = 17$)</td>
<td>42</td>
<td>83</td>
<td>67</td>
<td></td>
</tr>
</tbody>
</table>
Once again the prediction that babies viewing a familiar object should search more when hearing a novel label than a familiar label was confirmed for both 15-month-olds and 18-month-olds regardless of vocabulary level. The prediction that babies should search more when hearing a novel label relative to no label, however, was confirmed only for the younger babies, and only for the pre-spurt babies. As in Study 1, the weaker finding for the older babies is not caused by a drop in their searching in the Novel Label condition but by their higher baserate preference for looking at something else than the object before them in the No Label condition. It seems unlikely that 15 but not 18-month-olds use mutual exclusivity in this situation. A more plausible explanation is that 18-month-olds take advantage of the opportunity to explore their surroundings when the baseline condition does not make a specific demand of them.

When young children hear a novel label while attending to an object whose label is known, they tend to reject the term as a possible second label and instead search for another more appropriate referent. Even when only one candidate object is in view and when there is no obvious place to search, babies as young as 15 months spontaneously looked around the room as if searching for an appropriate referent for the novel word.

5. Discussion

By 2 1/2 years of age, children are highly consistent in mapping a novel object label to an object whose name is as yet unknown rather than to an object whose name is already known. When shown, for example, a spoon and a whisk and asked to hand the speaker the whisk, young preschoolers almost always hand over the whisk. This task was devised by Markman and Wachtel (1988) to reveal children’s use of mutual exclusivity. On their account, mutual exclusivity should lead children to reject a second label for an object with a known label and thus lead them to infer that the term applies to the object on display whose label is not yet known. Although this finding is highly replicable (e.g., Golinkoff et al., 1992; Merriman & Bowman, 1989; Merriman & Schuster, 1991), its interpretation and its significance has been called into question. This work addressed two of the main challenges about this role of mutual exclusivity.

The first concern raised about the earlier findings is that the children studied were old enough that they could have noticed that many terms in their vocabulary happened to be mutually exclusive and thus used mutual exclusivity as a convenient heuristic for acquiring some subsequent vocabulary. Yet mutual exclusivity and other word learning constraints have been postulated as being essential for helping children narrow down the otherwise intractably large hypothesis space for mapping words to concepts. The hypothesis tested here is whether mutual exclusivity is available to children before the naming explosion or vocabulary spurt which is seen on the average at 18–19 months. Vocabulary data were available for a majority of the children and the data were analyzed according to vocabulary level in addition to age.

Second is the question of whether children’s consistent performance in this indirect learning situation reflects mutual exclusivity per se or whether it can be
accounted for by children’s attempt to fill lexical gaps as Merriman and Bowman (1989) proposed. On their hypothesis, a novel object must be present to create a lexical gap and motivate the word learning. Mutual exclusivity instead can function even in the absence of an obvious candidate object and can motivate search for a referent that is out of sight.

A lexical gap is created when babies come across something they cannot label. If there is no unfamiliar object around, then no lexical gap is created. The key procedural change in these studies was to eliminate the novel object from the display in front of the babies to rule out the possibility that children could be attempting to fill a lexical gap. In Study 1 an opaque bucket and an object with a well-known label (e.g., a spoon) were placed on the table in front of the baby. The experimenter shook the bucket so the baby could hear that something was inside it, but could not see what it was. In Study 2 the familiar object was handed to the baby to play with. In Study 3 the bucket as an obvious location to search was removed.

Across the three studies, two kinds of measures and two kinds of comparisons generated several different predictions from mutual exclusivity. In Studies 1 and 2, the two measures were: (1) children’s selection of the visible object and (2) their tendency to search. We examined children’s selections of objects and search when they viewed, e.g., a spoon and a bucket on the table, and asked them either, “Can you show me a spoon?”, “Can you show me a gadget?”, or “Can you show me one?” The two comparisons looked at children’s selection or search in (1) the Novel Label condition relative to the Familiar Label condition, and (2) the Novel Label condition relative to the No Label condition.

The first prediction from mutual exclusivity is that if children are avoiding second labels for objects, they should indicate the spoon less when asked for a “gadget” than when asked for “a spoon.” In both Studies 1 and 2, young and old babies and pre- and post-spurt babies alike consistently avoided picking the spoon in the Novel Label condition relative to the Familiar label condition. This difference is predicted by mutual exclusivity and although it demonstrates children’s resistance to second labels, it also could be obtained if children preferred to see what was in the bucket and suppressed that tendency when asked for a well-known object. To evaluate this, we compared children’s tendency to pick, e.g., the spoon when asked for a “gadget” relative to when asked for “one.” In Study 1 both old and young babies selected the spoon less when asked for “a gadget” than when asked for “one,” but the effect was slightly stronger for the younger babies. Study 2 was designed to push the limits of children’s use of mutual exclusivity by handing them the toy to play with which made the toy more salient and interesting and thereby added strong inhibitory requirements to revealing mutual exclusivity. Although there was no interaction with age and both age groups combined indicated the spoon they were playing with less when asked for “a gadget” compared to “one,” the results did not reach significance for the age groups separately and appear stronger for the older children, who likely have more inhibitory control. We take this as reflecting developmental differences in the demands of the situation in Study 2 with inhibitory requirements affecting the results. So with this qualification, the selection data from Studies 1 and 2 support both predictions from mutual exclusivity.
The second measure across all studies was whether babies searched for an appropriate referent for the novel label. The argument here is that if babies have rejected a second label, they are then left with a word without a referent and, according to the principle of reference, they might be motivated to find an appropriate object. Again we evaluated babies’ search both when hearing a familiar label and when hearing no label. As predicted, in Study 1 babies at both ages and both vocabulary levels searched more when asked for “a gadget” than when asked for “a spoon.” In Study 2, while actually playing with the spoon, only the older babies searched more, but younger babies in this study hardly searched at all in any condition—they remained interested in the object in hand. Here the lexical constraints are pitted against this inhibitory requirement and younger babies do not interrupt their play to search for an appropriate referent. It is striking, however, that the older babies do. In Study 3, when the bucket was removed and there was no obvious place to search, babies at both age groups and both vocabulary levels searched more in the Novel label than in the Familiar label condition. So across all three studies this search prediction was confirmed for all babies at both ages and vocabulary levels, except for in Study 2 when the younger babies playing with a toy hardly search at all. The second prediction about babies’ search behavior is that babies should search more when asked for “a gadget” than when asked “for one.” In Study 2, as just mentioned, younger babies do not interrupt their play to search, but the older babies did and confirmed this prediction. In Studies 1 and 3 it is the younger babies that confirm the second prediction more strongly. This is particularly striking in Study 3 where younger babies spontaneously search the room despite there being an obvious object in front of them and no obvious location for another object. The reason younger babies show the effect more than older babies is that the baseline No Label condition revealed developmental differences in children’s exploratory behavior. When asked for “one” and thus free to do what they liked, younger babies were more likely to reach for the salient visible object whereas older babies were more interested in exploring the bucket or searching the room. In sum, although the effects were modulated somewhat by the exploratory and inhibitory demands of the tasks, across all three studies there is clear evidence that both younger and older babies search more when hearing a novel label than in the baseline No Label condition.

Taken together, the results rule out the lexical gap hypothesis and also contradict the findings of Mervis and Bertrand (1994) and Graham et al. (1998) who found that 1 1/2 year-old babies with limited vocabularies were not able to solve similar indirect word learning tasks. Their tasks most resemble the main selection prediction from mutual exclusivity that babies will avoid selecting the familiar object upon hearing a novel label relative to hearing a familiar label. Study 1 showed without any qualifications that babies from 15 to 19 months of age used this version of mutual exclusivity regardless of age and vocabulary level. One key difference between the studies is the number of objects that babies can choose as referents for the novel label. The current studies involved a two-choice task compared to the five-choice task of Mervis and Bertrand (1994) and three-choice task of Graham et al. (1998). This suggests that their tasks may have been too demanding for children, a notion that is
supported by the failure of Mervis and Bertrand’s babies to learn even from direct labeling with corrective feedback on the same task.

Another lexical constraint, the N3C principle, has been proposed to account for the findings from indirect word learning studies. The N3C principle is predicted to become available to children only at or beyond the time of the vocabulary spurt and it requires that a novel object be present at the time where a novel label is heard. Thus it would be hard for N3C to account for our findings. In Study 3 only a single familiar object was present and babies still searched. Golinkoff et al.’s framework specifies that children would either assume that the novel label is a second label for the familiar object and not search, or rely on pragmatic cues of the type posited by the social-pragmatic approach to provide a motivation for searching elsewhere (Mervis et al., 1994a).

The social-pragmatic view presents an alternative account of word learning to that of the lexical constraints approach. On that view, lexical principles are not needed for children to infer the referents of words. Instead, the findings from indirect word learning studies can be explained by children’s motivation to figure out the intended referent of the speaker and their expectation that the speaker will be a collaborative communicator. Children would reason that if the speaker had wanted the ball, he or she would have asked for the ball by using its conventional name (e.g., Bloom, 1998, 2000; Clark, 1990, 1997; Diesendruck & Markson, 2001; Gathercole, 1989; Tomasello, 2001). Although children use a range of sources of information to infer a speaker’s intentions to refer to an object, it is notable that indirect word learning studies render directly observable pragmatic cues such as eye gaze and pointing uninformative. Indeed, this ambiguity is a motivating force behind studies of mutual exclusivity. Thus, it is the child who must identify the relevant knowledge that will disambiguate the referent of the label and use this to choose among the possible referents. This knowledge may consist of lexical principles or knowledge of conventional communicative practices as well as other types of knowledge. As the current studies were not designed to distinguish between the mutual exclusivity and social-pragmatic accounts, both of these accounts remain theoretically possible. However, we know of no empirical studies from a social-pragmatic point of view that demonstrate that 15–18-month-old children are able to recruit knowledge of conventional communicative practices in cases where direct cues are uninformative, and we know of only one study with older children. Before turning to that study, we review the available evidence from studies that address young children’s abilities to infer the intended referent when a speaker uses a novel label.

Well before word learning beings, infants are able to attribute intention to entities that have features in common with animate beings (Johnson, 2000; Woodward, 1998, 1999; Woodward & Sommerville, 2000). Later, in the months preceding the vocabulary explosion, infants show understanding of the pragmatics involved when a speaker labels objects. Baldwin (1991, 1993; Baldwin, Markman, Bill, Desjardin, & Irwin, 1996) has explored whether 1–2-year-olds require some evidence of intent to refer to an object before learning a novel word. The alternative possibility would be that infants hear a label and associate it with whatever object that they are currently engaged with. In a series of “discrepant labeling” tasks, a speaker uttered a novel
label as the child was looking at a novel toy. The speaker, however, was looking into an opaque bucket and not at the visible toy (Baldwin, 1991). Even in the case where there was not a joint focus of attention, 18-month-olds learned that the word referred to the object that had been in the bucket and even 16-months-olds avoided making mapping errors. Thus before 18 months of age children are aware that conventional expressions of intent to refer must be present before mapping labels to novel objects.

Not all interaction with an object signals intent to refer. For example, adults may accidentally handle objects that they did not mean to refer to or may attempt to refer to objects that are not accessible. Older infants, between the ages of 18 and 24 months, are able to track speakers’ communicative intent in a range of situations that seem to require a more subtle understanding of indicators of intentional states. In a representative study, Tomasello and Barton (1994) presented 24-month-olds with a series of objects that were hidden in buckets. The speaker announced that her intention was to find “a toma,” yet she rejected the first two objects that she found with a scowl and then excitedly found the target object in a third bucket. In a control condition the speaker found the target on the first try. Children in both conditions learned the novel word for the target object, illustrating that children are aware that the achievement of a goal to refer was a condition for attaching the label to the target. Children at this age are also sensitive to frustrated goals when naming objects. In a similar study (Akhtar & Tomasello, 1996), 24-month-olds were shown a series of objects which were then hidden in different locations. The speaker announced that she wanted to find “a toma” and turned to a toy barn as her first choice of a location to search. However, the barn doors turned out to be locked. The experimenter expressed her disappointment and explained to the child that she was unable to open the doors. She then progressed to examine and handle objects in the other locations. In a control condition, the experimenter immediately found the target object behind unlocked doors. Follow-up comprehension tests showed that children learned the intended referent of the novel label equally well in both conditions, even when the referent remained locked in the barn. These findings and those of Tomasello and Barton (1994) have been replicated with 18-month-olds (Tomasello, Strosberg, & Akhtar, 1996).

The results of these studies convincingly demonstrate the abilities of very young children to infer communicative intent. Not only are children sensitive to pragmatic cues such as pointing and handling but are also able to update such cues by paying attention to expressions of disappointment or frustrated goals. It is notable, however, that none of the studies involve inferring the referent of a label under circumstances where the directly observable pragmatic cues are uninformative. On the contrary, for every object that the child might consider as a referent, the adult speaker gives a clear cue that she is now attending to it, and then marks whether to reject or accept the object for further consideration with explicit statements about an error, such as “whoops,” or emotional cues, such as excitement or disappointment, that signal whether a goal has been achieved or not. In contrast, in the indirect word learning studies possible referents are deliberately not marked by such action cues (see Baldwin & Tomasello, 1998, for a related argument).
As far as we know, only one study of children’s social-pragmatic abilities presents children with a task which resembles indirect word learning tasks in that the possible referents are not directly acted upon by the speaker in some manner. Akhtar, Carpenter, and Tomasello (1996) let 24-month-old children play with three novel objects in the company of two experimenters and one of the child’s parents. Each of the objects were singled out and described with excitement but never labeled. The parent and one experimenter then left the room and in their absence another novel object was introduced. All four objects were then placed in a row and the two adults returned and excitedly exclaimed: “Look, I see a gazzer!” The adults carefully avoided looking at any one particular object. In a control condition, the adults returned and said “Look at that!” with equal excitement. Children’s comprehension was later tested by asking the child to indicate the referent of the novel label. Only the children in the novel label condition correctly indicated the target object. In this study, rather than having each object singled out by the speaker, children themselves must identify that the relevant principle for choosing between the objects is the speaker’s emotional reaction to an object introduced in her absence. Although not exactly parallel to what would be required in the indirect word learning studies, this finding comes closest to demonstrating that young children may be able to recruit the relevant social-pragmatic knowledge in the absence of supporting cues. Note that the children in this study were 2 years old and that this has not yet been shown with 15–18-month-olds.

Suppose 15-month-olds are capable of such pragmatic reasoning. On the social-pragmatic account, babies in Studies 1 and 2 would reason that speakers would normally use the conventional term for an object to refer to it. Since the speaker used another term, she must have intended to refer to another object. As the babies know that the bucket may hide this other object, they search in the bucket. Study 3, however, presents a single, familiar object in full view on the table without any other objects around and more strongly suggests that the single object in view is a likely referent. Even if the speaker is not directly pointing to or looking at the object, the context alone is a strong pragmatic cue that the speaker intends to refer to that object. Thus, babies find themselves in a situation where the pragmatic cue to the speaker’s intention conflicts with the conclusion from his or her reasoning about communicative conventions like “if she had meant that, she would have said that.”

A stronger version of this type of conflict arises in Liittschwager and Markman’s (1994) Study 3 of second-label learning in 16-month-olds. Similar results with 18-month-olds were found in an unpublished study by Nowinski and Markman. In those studies, the speaker directly looks at and points to a familiar object while saying a novel label, yet the infants reject the object as a referent of the novel label. It is not clear how the social-pragmatic account explains a situation where all the directly observable pragmatic cues deviate from the conclusion that the infant would draw based on other sources of knowledge about a speaker’s intentions. Under what circumstances would a speaker point to one object while meaning to refer to another object, unless she also clearly indicated error, negative emotion, disappointment, or possibly deception? And, in turn, how would a child of 16 months have come to develop such strong assumptions about the existence of a communicative practice
where a speaker points to a different object than the one she intends to label without any of the above markers? Even allowing for conflicts between pragmatic cues and other sources of knowledge, it seems unlikely that 16-month-old babies’ reasoning that “if she had meant that, she would have said that” alone would be a strong enough motivator to override the pragmatic cues of pointing and eye gaze and cause babies to search elsewhere. The social-pragmatic explanation specifically does not include a mutual exclusivity assumption so there is nothing in principle to prevent the baby from assuming that the speaker might use two labels to refer to the same object. The social-pragmatic explanation of why 16-month-olds reject second labels for objects is a complicated story whereby “if she meant that, she would have said that” outweighs the fact that the speaker is obviously labeling the object. This seems much less compelling than the straightforward explanation based on the mutual exclusivity assumption where children would reason “that can’t be a crimp, it’s a spoon.”

Can the success of 15-month-olds on these indirect word learning tasks be explained by social-pragmatic reasoning? Maybe—the jury is still out. But at this point, related evidence supports the claim that mutual exclusivity is necessary to account for our findings because: (1) there is explicit evidence that 16-month-olds avoid second labels for objects as predicted by mutual exclusivity even in the face of clear, unambiguous pragmatic cues that the speaker did intend to refer to the given object with the second label, and (2) there is a lack of evidence that 15-month-olds have the ability to do the necessary reasoning when explicit social-pragmatic cues are absent or ambiguous.

In conclusion, our findings demonstrate both that mutual exclusivity is available to children before the age of the vocabulary spurt and one of the purposes it serves is to promote some kinds of indirect word learning. The results highlight the way in which lexical constraints can serve as a resource for children when directly observable cues from the speaker or objects in the environment are lacking. By 15 months of age babies reject second labels for objects and actively search for appropriate referents of a word rather than always relying on a speaker to direct their attention to a given referent or on having the mapping triggered by viewing a novel object as a candidate referent. Mutual exclusivity and the principle of reference provide children with adequate grounds to actively search for appropriate referents of novel terms. In this way, lexical constraints add a layer of resilience to babies’ word learning, enabling them to form word-referent mappings even when speakers are not clear and referents are not visible.

The findings are also significant in demonstrating that word learning constraints play a role in the fast acquisition of words seen at around 19 months of age. Our findings do not, of course, address questions about the ultimate origins of this bias (see Woodward & Markman, 1998). Whether and what kinds of exposure to linguistic input are relevant to working out this assumption remains an open question. Our findings do address, however, the question of whether word learning constraints are necessary to account for young children’s ability to rapidly acquire vocabulary or whether they are merely dispensable heuristics that children discover after having already become capable of figuring out the meanings of novel terms in some other way. When the ability to recruit mutual exclusivity in this indirect word learning
situation was only found in 2-year-olds it was a fair criticism to argue that children might be using a learned heuristic that developed as a consequence of having gone through the naming explosion (MacWhinney, 1989). But the success of 15-month-olds in the present study rules out this explanation. Instead of being a late developing learned heuristic, mutual exclusivity is available to serve a vital role in the rapid learning of novel words for babies at the time of the naming explosion.

Acknowledgments

We are grateful to the many parents who generously volunteered to have their babies participate in these studies and to Bing Nursery School for providing research facilities. We thank Colleen M. Fitzsimmons for her help in coding the data and the reviewers for their constructive suggestions. This work was supported in part by NSF Grant BNS-91092 to Ellen M. Markman.

References


