



Published in final edited form as:

*J Exp Child Psychol.* 2009 February ; 102(2): 182–195. doi:10.1016/j.jecp.2008.04.006.

## Compliance, Conversion, and Category Induction

Vikram K. Jaswal, Olivia K. Lima, and Jenna E. Small

*University of Virginia*

### Abstract

When children hear an object referred to with a label that is moderately discrepant from its appearance, they frequently make inferences about that object consistent with the label rather than its appearance. We asked whether 3-year-olds actually believe these unexpected labels (“conversion”), or whether their inferences simply reflect a desire to comply with the considerable experimental demands of the induction task (“compliance”). Specifically, we asked how likely children would be to pass an unexpected label on to another person who had not been present during the labeling event. Results showed that children who used an unexpected label as the basis for inference passed that label on to another person about as often as they could remember it. This suggests that children’s label-based inferences do reflect conversion rather than mere compliance.

---

From an early age, children are highly receptive to testimony as a source of information. Among their first 50 words, for example, many have an idiosyncratic term they use specifically to request information from other people (e.g., “tha?”; Nelson, 1973), and by 2.5 years of age, they make frequent use of “What’s that?” questions to request the names of unfamiliar objects (Clark, 1991) or their functions (Kemler Nelson, Egan, & Holt, 2004). When testimony concerns something about which children have no expectations, such as the names or functions of novel objects, their deference to that testimony may not seem particularly surprising. But in some cases—specifically in studies on category induction—young children defer to testimony even when it conflicts with their expectations (e.g., Gelman & Markman, 1986; Graham, Kilbreath, & Welder, 2004; Jaswal, 2004; Sloutsky & Fisher, 2004).

For example, in Gelman and Markman’s (1986) classic study, preschoolers were shown a drawing of a fish and one of a dolphin. The experimenter explained that the fish “stays underwater to breathe,” whereas the dolphin “pops above the water to breathe.” Then, children were shown a drawing of an animal that looked like a dolphin, but which the experimenter unexpectedly referred to as a “fish,” and they were asked how it breathed. Children resolved this conflict by favoring the label: Even though the animal looked like a dolphin, they inferred that it breathed underwater, like the fish.

The deference to unexpected labels observed in studies on category induction is somewhat surprising because there are other situations in which very young children resist or reject testimony that conflicts with their expectations. For example, Pea (1982) found that by 24 months of age, most children spontaneously objected when a speaker referred to a car as a “ball.” Similarly, Koenig and Echols (2003, Study 1) found that nearly all 16-month-olds who

---

Correspondence concerning this article should be addressed to Vikram Jaswal, Department of Psychology, 102 Gilmer Hall, P.O. Box 400400, University of Virginia, Charlottesville, VA 22904-4400; email: jaswal@virginia.edu. Phone: (434) 982-4709; Fax: (434) 982-4766.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

heard an experimenter mislabel a series of objects (e.g., refer to a ball as a “duck”) responded by providing the veridical label at least once. Koenig, Clément, and Harris (2004) showed that preschoolers discounted information from a speaker who mislabeled familiar objects: They were less likely to trust new information from a speaker who had mislabeled familiar objects than from one who labeled them correctly (see also Jaswal & Neely, 2006).

One reason children in studies on category induction may be more deferential to unexpected labels than those in the studies just described has to do with the nature of the labeled stimuli. In many studies where children seem to reject a label, that label and children’s expectations are highly discrepant. For example, in Koenig and Echols (2003), the experimenter mislabeled pictures of typical exemplars from familiar categories, sometimes even using labels that crossed an ontological boundary (e.g., a chair was referred to as a “cat”). In contrast, most category induction studies use stimuli that have been carefully selected so that the labels the experimenter uses, while unexpected, are still plausible. For example, Gelman and Markman (1986) used anomalies actually found in nature: One stimulus set involved a kaibab, a squirrel that has unusually long ears. Although a kaibab looks somewhat like a rabbit, it also has some squirrel-like features (e.g., a long tail) and so could possibly be a squirrel (see Gelman & Coley, 1990). Although they did not frame their research in terms of children’s deference to an unexpected label, Sloutsky and Fisher (2004) demonstrated experimentally the role that the plausibility of the label plays: They found in several studies that children were more likely to make a label-based inference about an item the more perceptually similar the item was to the labeled category (see also Graham, Kilbreath, & Welder, 2004).

Another reason that children in studies on category induction may seem relatively deferential—and the focus of the study presented here—has to do with the demand characteristics of the category induction task. In both category induction studies and mislabeling studies, an adult experimenter names an object using a label that conflicts with the object’s appearance (e.g., a chair is called a “cat” or an animal that looks like a dolphin is called a “fish”). In mislabeling studies, the children’s spontaneous reactions to this mislabeling event are observed (e.g., Pea, 1982; Koenig & Echols, 2003), or the way in which they treat subsequent information from the inaccurate versus an accurate speaker is measured (Koenig et al., 2004).

In contrast, in studies on category induction, an adult refers to an item with a label that does not match children’s expectations, and then asks children to make either a label-based or appearance-based inference. Children’s spontaneous reactions to the unexpected labels are not of interest (and they are rarely reported; see Jaswal, 2004), and there is no competing testimony that is consistent with their expectations to which children can turn. Because speakers normally provide information that is relevant to the topic at hand (Grice, 1975), a reasonable pragmatic assumption in the category induction task is that the adult believes that the unexpected label is relevant to the inference they are asking the children to make. As a result, children may make a label-based inference because they think this is what the experimenter wants them to do rather than because they believe that the labeled object is a member of the named category (e.g., Aronsson & Hundeide, 2002; Donaldson, 1978; Siegal, 1997; Siegal & Surian, 2004).

Indeed, Siegal, Waters, and Dinwiddy (1988, Experiment 2) found that 4- to 6-year-olds recognize that people sometimes make responses to satisfy experimental demands rather than because they actually believe in what they are doing. In that study, children watched a video clip of a puppet performing a conservation of number task involving two rows of buttons. The puppet initially indicated that one row had more buttons than the other, but when the experimenter changed the spatial configuration of the two rows and asked a second time which row had more buttons, the puppet selected the other row. Interestingly, most children indicated that the puppet changed his response “just to please the grownup,” not because he “really thought [it] was true.” In contrast, when the puppet indicated the same row of buttons both

before and after the transformation, most children indicated that the puppet's response reflected his true belief rather than a desire to please the grownup.

The distinction we wish to make is captured by the terms *compliance* and *conversion* (MacDonald, Nail & Levy, 2004; Mugny, 1984; Nail, MacDonald & Levy, 2000). When confronted with a message that conflicts with their own beliefs, children (and adults) may behave in a manner consistent with the message for social reasons ("compliance") rather than because their beliefs have actually changed ("conversion"). To our knowledge, no one has yet considered the extent to which children's performance on a category induction task reflects compliance rather than conversion.

This is a crucial question because the category induction task has been used by a number of researchers interested in the structure of children's categories (e.g., Gelman, 2003; Jaswal, 2004; Sloutsky & Fisher, 2004). We will not describe the on-going controversy about whether children's categories are theory-based or similarity-based because the present work was not designed to speak directly to this debate (but see Sloutsky, Kloos, & Fisher, 2007a, 2007b; Gelman & Waxman, 2007). Rather, our goal was to develop a new procedure that could be used to investigate whether children who use an unexpected label as the basis for inference actually believe that label.

In our new procedure, one experimenter (E1) provided an unexpected label and asked children to make an inference. She then left the room with the excuse that she had to get something; she was replaced by a second, ostensibly naïve experimenter (E2). This second experimenter engaged the children in casual conversation while they waited for E1 to return. Of interest was whether children who used E1's unexpected label as the basis for inference about an object would pass that label on to E2 when she casually asked them about the name of that object.

Our use of two experimenters was motivated by several studies showing that young children can use someone's presence or absence to make inferences about that individual's likely knowledge (e.g., Akhtar, Carpenter, & Tomasello, 1996; Diesendruck & Markson, 2001; O'Neill, 1996; Tomasello & Haberl, 2003). For example, Diesendruck and Markson (2001) showed that 3-year-olds did not expect that a puppet would know a novel fact (e.g., which of two objects had been given to the experimenter by her uncle) unless that puppet had been present earlier when the experimenter stated that fact. Children did, however, expect that the puppet would know the name of a novel object—even if the puppet had not been present during the labeling event. Based on these previous studies, we reasoned that 1) children would not assume that E2 knew what had happened when she was out of the room, but 2) they would assume that she knew the conventional names of objects. Thus, when E2 asked about the name of an object, children would not feel pressure to provide the unexpected label E1 used earlier unless they actually believed that to be its conventional label.

Suppose, for example, that E1 refers to a key-like object as a "spoon," and children make a label-based inference, inferring that it is used to eat cereal rather than start a car. If children later tell E2 that the key-like object is a "spoon," this would suggest that they actually believe it is a spoon. If, however, they tell her that it is a "key," this could suggest that their earlier spoon-like inference had been due to compliance.

It is also possible, however, that children will believe E1's assertion that the key-like object is a "spoon," but that they will later refer to it as a "key" because of memory limitations. After all, the label the experimenter uses is counter-intuitive and may therefore be difficult to remember. Traditionally, studies on category induction have focused exclusively on the inferences children make immediately on hearing an unexpected label, not on their ability to retain that label. In our own previous work using the category induction procedure (Jaswal, 2004, 2007; Jaswal & Malone, 2007; Jaswal & Markman, 2007), we have informally asked

children to name the objects after completing the study, and have found that many who seemed deferential to the unexpected labels earlier nonetheless reverted back to naming the objects according to their appearance.

Thus, another group of children participated in a memory control condition. In this condition, children took part in the category induction task, just as those in the experimental condition did. However, rather than having E2 later ask about the names of the objects, E1 later asked children what she had called them. We expected that children in both conditions would be equally likely to make label-based inferences—the two conditions do not differ at this point in the procedure. The crucial comparison is the difference between the number of labels children in the memory control could remember and the number children in the experimental condition passed on to E2. If those in the memory control remember more than those in the experimental condition pass on, this would suggest that at least some of the label-based inferences children in the experimental condition made were due to compliance.

We chose to study 3-year-olds because this age is the focus of a number of previous studies using the category induction procedure (e.g., Gelman & Markman, 1987; Jaswal, 2004; Jaswal & Malone, 2007), and, as noted earlier, various theoretical claims have been made on the basis of the performance of children this age (and younger) on this task (e.g., Gelman, 2003). It is important to note that there is already evidence showing that the category induction procedure does not always result in conversion. In Jaswal (2004), for example, 3-year-olds ignored the unexpected labels the experimenter used on about one-third of the trials, often expressing their skepticism explicitly (e.g., “That doesn’t look like a spoon,” or “That’s not a spoon; it’s a key!”). Clearly, when children ignore or verbally reject an unexpected label, this reflects neither compliance nor conversion. Our primary interest is therefore in those cases where children do defer to the unexpected label the experimenter uses: Does this deference reflect a change in children’s beliefs about what the labeled object actually is? Or does it reflect mere compliance?

## Preliminary Study

We first obtained a baseline measure of children’s expectations about the stimuli to be used in the experimental induction task. We wanted to ensure that when the items we used were referred to neutrally, without labels, children would consistently base their inferences on their appearance. Further, we wanted to ensure that under these conditions, they would later describe those items to another person using labels consistent with their appearance.

## Method

**Participants**—Eight 3-year-old children ( $M$  age: 3 years, 6 months; range: 3;2 to 3;11; 4 boys) participated in a single laboratory visit. The children in this study, and the main study, were recruited through a database of local families who had expressed an interest in participating in research. Most children were Caucasian and from middle- or upper-middle-class backgrounds, and all were fluent English speakers.

**Stimuli**—Four sets of stimuli were prepared from a digital library of photo-objects (Hemera Technologies, Canada). Photos representing prototypical exemplars of eight familiar categories were selected and arranged into pairs to form four stimulus sets: key-spoon, shoe-car, toothbrush-pen, and hat-cup. These stimuli served as “standards.” For each stimulus set, two additional “typical” exemplars were created by modifying the color and/or size of the standards. Finally, a “hybrid” exemplar was created from each set, which looked mostly like one of the standard exemplars of that set, but which shared some features of the other. For example, the key-like hybrid looked mostly like a key, but it had some spoon-like features (e.g., a silver, wide reflective base). In addition to the key-like hybrid, there was a shoe-like hybrid (with some features of a car), a hat-like hybrid (with some features of a cup), and a

toothbrush-like hybrid (with some features of a pen). Figure 1 shows the four hybrid items. It is important to note that these stimuli were not ambiguous in appearance. Although they had features of both categories, the Preliminary Study was designed to confirm that children would have a robust expectation that these hybrids were members of the category they most resembled.

Each stimulus was approximately 2–4 inches in width, and 2–4 inches in height. Each one (and its left-right reverse) was printed in color and mounted onto a foam core base, which allowed it to stand independently.

Eight additional color photographs were selected to represent a scene or object with which the function of each stimulus could be demonstrated. For example, a photograph of a car was used to demonstrate that a key could be used to start a car, and a photograph of a bowl was used to demonstrate that a spoon could be used to eat cereal. These photos were approximately 6 x 6 inches, and were printed and mounted onto 8.5 x 11-inch easels. Table 1 shows the complete list of functions for each stimulus set.

Finally, a set of warm-up stimuli was constructed, consisting of exemplars of dolls and shovels, and background photographs of a bed and a bucket. As will be described below, the doll was put to sleep in the bed, and the shovel was used to pretend to scoop sand into the bucket.

**Procedure**—On arriving at the lab, children were greeted by two female experimenters (E1 and E2). These two experimenters played with the children in the waiting room while the parent completed the consent forms (usually for around 10 min). We chose to have both experimenters interact with the children before the session because we wanted to be sure that they were familiar and that children were comfortable with both of them. This was especially important because children later had to provide verbal responses to questions posed by E2. In pilot work, we found that 3-year-olds were extremely reticent when they were not familiar with E2. When the parent had completed the consent forms and a child appeared comfortable, E1 said goodbye to E2, and led the child to the experimental room.

E1 began the session with a warm-up trial to familiarize children with the procedure. First, she showed the background photo of a bed. She explained aloud and demonstrated that the doll stimulus could sleep in the bed. Next, she showed the background photo of the bucket, and explained and used the shovel stimulus to pretend to scoop sand into the bucket. Finally, E1 presented the child with another doll and shovel, one at a time, and asked where each went. Correct selections were praised, and errors were corrected. Four children had to be corrected and were asked to make inferences about an additional doll and shovel.

After children had successfully placed a doll with the bed and a shovel with the bucket, E1 removed the easels from the table, revealing the exemplars of the dolls and shovels that had been used during that warm-up trial. She explained that she needed to go to a different room to get the next set of stimuli, but that she would call E2 to wait with the child. She opened the door and called for E2, who appeared and agreed to stay. E1 left the room and closed the door behind her.

Because children had interacted with both experimenters in the waiting room, children might infer that E2 somehow knew what had happened in the room even though she had not been there. In an attempt to prevent children from coming to this conclusion, each time E2 was asked to wait with the child, she seemed surprised and ignorant about the game the children had been playing and the items sitting on the table. She casually asked about the names of each one (“Wow! Look at these! These are neat pictures! What’s this one?”). All children used the correct labels (“doll” and “shovel”) when referring to the warm-up stimuli. Shortly after the child had

provided a name for each item, E1 returned, carrying a new set of stimuli, and E2 left, closing the door behind her. E1 then began the test trials.

Test trials resembled the warm-up trial. For each of the four stimulus sets, children watched as E1 demonstrated and explained that each standard item could be used for a particular function. For example, E1 showed and explained that a key could be used to start a car, and a spoon could be used to eat cereal from a bowl. Following the demonstration, children were presented with the three test items for that stimulus set (two typical and one hybrid), one at a time, in a pseudo-random order such that the hybrid item was presented first, second, and third at least once each across the four trial blocks. E1 referred to each of the test items neutrally, without labels: “Look at this! Can you show me where this one goes?” Hybrid items were introduced in exactly the same way as the typical ones—that is, no special attention was drawn to them. Children received neutral feedback following a response (“Okay!”).

After children responded to the three test items of a given set, E1 called E2 to wait with the child while she went to get the next set of stimuli. As in the warm-up trial, E2 expressed an interest in the items sitting on the table, casually asking about the names of each one. The order in which E2 asked about the names was the same as the order in which children had been asked by E1 to make an inference about each one. Thus, they were asked about the hybrid item first, second, and third at least once across the four trial blocks. When all the stimuli had been named, E1 returned with the next set of stimuli, and E2 left. E1 then repeated the procedure using the new stimulus set. The four sets of stimuli were presented in one of four orders, according to a Latin Square design.

Coding in this study and in the Primary Study was conducted from videotape, and consisted of noting which inference the child made about each test item, and what she or he called them when asked about their names. A second coder coded a random selection of 25% of sessions from the Preliminary Study and the Primary Study. Reliability was excellent, with 100% agreement on children’s selections during the inference task and 98% agreement on their verbal responses during the naming one.

## Results and Discussion

Children made appearance-based inferences about the typical and hybrid items on 98% and 100% of trials, respectively. When E2 later asked about the names of those items, children responded by providing labels that matched their appearances on 97% of the trials for the typical items and 100% for the hybrids. For example, children used both the typical key and the key-like object to start the car and later called both “keys,” and they used the typical spoon to eat cereal from the bowl and called it a “spoon.”

From these results, we can be sure that the hybrid objects were seen as members of the categories they were designed to resemble. Even though the hybrids had features of two categories, children did not see them as ambiguous blends of the two. They always made inferences about the hybrids consistent with the particular category they were designed to resemble most, and always labeled them using the label of the particular category they were designed to resemble most. Clearly, children had robust expectations about the category to which each hybrid belonged.

## Primary Study

The Primary Study was designed to investigate whether children’s label-based inferences reflect compliance or conversion. The procedure was similar to that used in the Preliminary Study, except E1 referred to each hybrid item using a label that conflicted with the one children in the Preliminary Study had always used (e.g., E1 referred to the key-like object as a “spoon;”

children in the Preliminary Study had always referred to it as a “key”). Previous work with a similar set of stimuli has shown that under these conditions, 3-year-olds often base their inferences on the unexpected labels (Jaswal, 2004; Jaswal & Malone, 2007).

In the experimental condition, E1 then left the room and E2 entered, and casually asked children what each object was called. In the memory control, E1 also left the room, but she returned immediately and asked children what she had called each object. Our interest was in whether some of children’s label-based inferences reflected compliance due to the experimental demands of the task. If so, children in the memory control should remember more unexpected labels than those in the experimental condition pass on to E2.

We also included a trial at the end of the study in order to ensure that our procedure could identify a case of compliance. As noted in the Introduction, previous category induction studies have used labels that are only moderately discrepant from children’s expectations. On this final trial, we purposely violated this convention by using a label that crossed the ontological boundary between natural kinds and artifacts: The experimenter referred to a picture of a typical horse using the name of an artifact (e.g., “key”). Children had to choose between making an inference consistent with the label, but which was not appropriate for a horse (e.g., using it to start a car), or an inference that was neither consistent with the label nor appropriate for a horse (e.g., using it to eat cereal). We expected that some 3-year-olds would make a label-based inference simply to satisfy the experimental demands of the task. But we doubted that they actually would believe that the horse was, for example, a key. Thus, we expected that children in the memory control would remember that E1 called the horse a “key.” But we did not expect that children in the experimental condition would tell E2 that the horse was a “key.”

## Method

**Participants**—Thirty-two 3-year-olds (mean: 3 years, 6 months; range: 3;0 to 4;0; 16 boys) participated. None had participated in the Preliminary Study. One additional child was excluded because of parental interference.

**Stimuli and Procedure**—The stimuli and procedure were similar to those used in the Preliminary Study. As in the Preliminary Study, there were four trial blocks, each consisting of three inference test trials (two involving typical items and one involving a hybrid). In the Primary Study, however, E1 introduced each test item by referring to it twice by name. She referred to the typical test items using labels that matched their appearance, and to the hybrid test items with labels that did not (e.g., the key-like object was referred to as a “spoon;” see Figure 1). As in the Preliminary Study, children were asked to make an inference about each test item. At the end of each stimulus set, E1 left the room, claiming that she needed to get the next set of stimuli.

Children participated in one of two conditions. In the experimental condition ( $n = 16$ ), E2 entered and asked children what each item in that set was called. In the memory control condition ( $n = 16$ ), E1 returned and asked children, “What did I call this?” for each item in that set. The average time between the conclusion of the inference task and the start of the naming task in both conditions was about 15 s.

After the final inference of the fourth trial block in both the experimental and memory control conditions, E1 held up a picture of a typical exemplar of a horse (3 x 3 inches), referred to it twice either as a hat, key, car, or toothbrush, and asked children to make an inference about it. Clearly, none of these labels is appropriate for a horse. The particular label E1 used and the particular inference choices available depended on which stimulus set had comprised the fourth trial block. For example, if the fourth trial block had involved keys and spoons, the background photo showing a van and the one showing a bowl of cereal remained on the table. Children

heard E1 refer to the horse as a “key,” and they had to decide whether it was used to eat cereal or to start the car. E1 then left the room and either was replaced by E2, who asked about the names of all of the objects from the fourth trial block and the horse (in the experimental condition), or E1 returned to ask what she had called each object (in the memory control).

## Results and Discussion

**Horse trial**—We describe results from the horse trial first, because it was designed to show that children might make a label-based inference because of compliance, and also to ensure that our procedure could identify a case of compliance. Because this involved a single trial, we will report results in terms of the number of children who responded in a particular way. Seven of the 16 children in the experimental condition made a label-based inference even though E1 used a label that crossed an ontological boundary (e.g., referred to a horse as a “key”); one made the opposite inference, and eight simply refused to respond. Importantly, however, no child used the inappropriate label when E2 entered and asked what the horse was called; all 16 referred to it as a “horse.”

The memory control showed that this was not because they had forgotten the label E1 used. In this condition, 10 of 15 children made a label-based inference about the horse (one child did not receive this trial due to experimenter error). When E1 left the room and then returned to ask what she had called it earlier, 7 of those 10 children remembered, saying that she had called it, for example, a “key.” The other three misremembered, claiming that she had called it a “horse.”

The results of this horse trial are important because they demonstrate that the pragmatic demands of the induction task can lead children to make a label-based inference even when the label is clearly inappropriate. Although some children objected and refused to respond when the experimenter mislabeled the horse, about half went along with the inappropriate label she used. The results of this trial are also important because they demonstrate that our procedure can identify such cases of compliance: Children in the memory control tended to remember the bizarre label E1 used, but those in the experimental condition never passed that label on to E2.

**Typical items**—Each participant received eight trials involving typical items, and so results will be reported in terms of the average percentage of trials on which children responded in a particular way. As in the Preliminary Study, in both the experimental and memory control conditions, children nearly always inferred that the typical items had functions consistent with their appearance (on at least 99% of trials). Those in the experimental condition later described them to E2 using labels consistent with their appearance on 99% of trials. Those in the memory control condition later reported to E1 that she had called them by labels consistent with their appearance on 94% of trials.

**Hybrid items**—Each participant received four trials involving hybrid items, and so results will be reported in terms of the average percentage of trials on which children responded in a particular way. Recall that children heard each of the four hybrids referred to with labels that did not match their appearance. As Figure 2 shows, most inferences in both conditions were consistent with those labels rather than with the appearance of the hybrid: Children made label-based inferences about the hybrids on an average of 53% (SD = 33%) and 67% (SD = 33%) of the trials in the experimental and memory control conditions, respectively. Although children were slightly more likely to make label-based inferences in the memory control than the experimental condition, the two conditions did not differ from each other,  $t(30) = 1.22$ ,  $p = .23$ . Consistent with a good deal of previous research using the category induction task, then, children in both conditions were often initially deferential to the unexpected labels they heard



the experimenter use, even though those labels conflicted with what we knew from the Preliminary Study their expectations to be (e.g., Gelman & Markman, 1986, 1987; Jaswal, 2004; Sloutsky & Fisher, 2004).

The critical and novel part of our procedure was the naming task. In the experimental condition, an ostensibly naive E2 replaced E1 and casually asked children what the hybrid from a particular set was called. In the memory control condition, E1 asked explicitly, “What did I call this?” We describe results for naming trials that followed appearance-based inferences first, and then turn to naming trials that followed label-based inferences.

As Figure 2 shows, when children ignored the unexpected label E1 used and made an appearance-based inference, they almost always later referred to the hybrid with the label consistent with its appearance. They did so 100% of the time in the experimental condition, and 95% of the time in the memory control. Consider, for example, a trial on which E1 referred to a key-like object as a “spoon.” Children in the experimental condition who ignored E1’s unexpected label and used the object to start a car rather than eat cereal always told E2 that it was a “key.” Interestingly, children in the memory control condition who ignored E1’s unexpected label later mistakenly claimed that she had called it a “key.” We will return to this intriguing error in the General Discussion.

Our primary interest was in how children would name the hybrids when they had earlier made a label-based inference. Because the memory control provides a baseline of how many of the unexpected labels children could actually remember, we describe results from this condition first. As Figure 2 shows, on those trials when children in the memory control made a label-based inference, they subsequently recalled the unexpected label E1 had used, on average, 52% of the time. The other 48% of the time, they incorrectly reported that she had referred to the hybrid using a label consistent with its appearance. Given that children in the memory control could remember only half of the unexpected labels E1 used, we would not expect those in the experimental condition to pass on to another person *more* than 50%. However, if they pass on *fewer*, this would suggest that some of their initial label-based inferences were due to compliance rather than conversion.

As Figure 2 shows, on those trials when children in the experimental condition made a label-based inference, they subsequently passed the unexpected label on to E2, on average, 39% of the time, used a compound noun that combined the label and appearance (e.g., “shoe-car”) 9% of the time, and used a label consistent with the hybrid’s appearance the remaining 52% of the time. Although children were slightly less likely to report the counter-intuitive label in the experimental condition than in the memory control (39% vs. 52%), this difference was not significant,  $t(27) = 1.03$ ,  $p = .31$ , Cohen’s  $d = .38$ .<sup>1</sup> A power analysis indicated that it would take over 107 participants in each condition for an effect of this size to become significant, and even then, the magnitude of the difference between the two conditions (13%) would appear to be quite small.

In short, children in the experimental condition passed on about as many unexpected labels to E2 as those in the memory control condition could remember. This suggests that when 3-year-olds made a label-based inference about a hybrid, they usually did so because they actually believed that label, not simply because of the experimental demands of the situation.

**Individual response patterns**—One important question concerns the extent to which the group data in each condition are representative of individual response patterns. Table 2 shows

---

<sup>1</sup>Note that the degrees of freedom for this t-test depended on the number of children who made a label-based inference. Because 14 children in the experimental condition and 15 children in the memory control did so, the df for this test was 27.

the number of children reporting 0, 1, 2, or 3 of the four unexpected labels during the naming task to E2 (in the experimental condition) or E1 (in the memory control) as a function of the number of label-based inferences they made initially. There are two important features to note from the table. First, during the initial inference task, most children made label-based inferences on at least half of the trials: 11 out of 16 in the experimental condition, and 14 out of 16 in the memory control did so. Additionally, the distribution of the number of children making each number of label-based inferences was similar across the conditions,  $\chi^2(4, N = 32) = 2.33, p = .67$ . Thus, most children were neither completely credulous nor completely skeptical initially, consistent with the findings from other category induction studies with 3-year-olds (Gelman & Markman, 1987; Jaswal, 2004).

Second, 10 of the 11 children (91%) in the experimental condition who made two or more label-based inferences passed on at least one of those unexpected labels to E2. Similarly, 12 of the 14 children (86%) in the memory control who made two or more label-based inferences recalled at least one of those labels later. Thus, results were not driven by a small subset of children who were both willing to believe anything E1 said and who had excellent memories.

## General Discussion

There are at least two reasons why children participating in a category induction study might make an inference consistent with an unexpected label they just heard: They might actually believe that label, or they might simply feel pressure to comply with the label because of the considerable experimental demands of the task. The study reported here is important because it is the first to acknowledge and explicitly address these possibilities. Results suggested that 3-year-olds' label-based inferences about the hybrid stimuli reflected conversion rather than compliance. Children in the experimental condition passed on about as many of the unexpected labels for the hybrid items to another, ostensibly naive person as children in the memory control could remember.

Our "horse trial" demonstrated that the demands of the induction task could lead children to make a label-based inference, and that our procedure could identify a case of compliance: In this trial, the experimenter referred to a horse with the label of an inanimate object, violating the convention used in previous studies of category induction, where the unexpected label has typically been from the same ontological category as the named object. Nevertheless, about half of the children made a label-based inference about the horse. Of these children, most in the memory control could later remember the inappropriate label the experimenter had used, but none in the experimental condition transmitted it to another person. Thus, the initial label-based inferences made about the horse on this trial most likely reflected compliance due to the experimental demands of the induction task.

That some children were deferential on the horse trial is consistent with other research showing that children sometimes respond in particular ways in order to "please the grown-up" (e.g., Siegal et al., 1988). It is also consistent with research from the child eyewitness testimony literature, which has demonstrated that very young children may respond in a particular way because they have as their primary goal a desire to affiliate with the adult interlocutor (e.g., Ceci & Bruck, 1993). This may conflict with the experimenter's goal, which is of course to obtain responses that reflect the children's true beliefs (Aronson & Hinde, 2002; Siegal & Surian, 2004).

In contrast to children's label-based inferences on the horse trial, their label-based inferences about the hybrids seemed to reflect conversion. Children in the experimental condition were willing to pass on a number of unexpected labels to a second experimenter who had not been present during the labeling episode. We reasoned that children would not report those labels

to E2 unless they actually believed them to be the conventional terms for those objects (e.g., Diesendruck & Markson, 2001). One objection could be that children did not see E1 and E2 as independent. If children thought E2 somehow knew the unexpected labels E1 had used, they might have felt the same pressure to comply even though the naming task was conducted by E2. We think this is unlikely for three reasons.

First, the same procedure was used on the trial involving the horse, and children in the experimental condition always told E2 that it was a “horse” rather than the unusual label E1 had used earlier. If children had felt pressure to provide E2 with the unusual label E1 had used, we would have expected them to report that label. Second, we attempted to make it clear that E2 was not familiar with the stimuli: She feigned surprise each time she entered the room and saw a new set of objects (“Wow! Look at these!”). Finally, as noted in the Introduction, a number of studies have shown that very young children readily make use of a person’s presence or absence to make inferences about their likely knowledge (e.g., Akhtar, Carpenter, & Tomasello, 1996; Diesendruck & Markson, 2001; O’Neill, 1996; Tomasello & Haberl, 2003).

One question concerns why children seemed to show evidence of compliance on the horse trial, and conversion on the hybrid trials. This difference likely stems from the fact that the label E1 provided during the horse trial was highly discrepant from children’s expectations—it crossed the natural kind-artifact ontological boundary. Although many children were willing to play along with the experimenter when she referred to the horse as a “key,” for example, they did not actually believe that it was a key. In contrast, the labels E1 provided during the hybrid trials were unexpected, but plausible: They named other artifacts that were roughly the same size and shape as the depicted artifact.

In most studies on category induction, including the one reported here, children are generally deferential to the moderately discrepant labels they hear an adult use. There are a number of other situations in which children also defer to an adult’s label rather than favoring their own. For example, even if they over-extend the word “dog” to all four-legged animals, they eventually do learn to refer to non-dogs appropriately (e.g., Mervis, Pani, & Pani, 2003). Similarly, children sometimes coin new words to fill gaps in their lexicon (e.g., “don’t broom my mess” for “don’t sweep my mess”), but these innovations fall out as soon as children learn the appropriate vocabulary (Clark, 1982, 1991).

This raises an intriguing question: Why should children weight what someone else says more heavily than their own expectations? We suspect that this deference reflects the fact that children recognize from an early age that language is conventional (e.g., Clark, 1991; Henderson & Graham, 2005), and that they have a general inclination to trust what adults tell them (e.g., Csibra & Gergely, 2006; Dawkins, 2004; Harris, 2002; Reid, 1764/1997). Despite this general inclination to trust adults, children are not entirely credulous: They may respond skeptically if an adult says something that is highly discrepant with their expectations (e.g., Koenig & Echols, 2003; Pea, 1982), if an adult expresses uncertainty about something that is moderately discrepant (Jaswal & Malone, 2007), or if an adult has been wrong in the past (Jaswal & Neely, 2006).

The memory control condition showed that children had difficulty remembering the unexpected labels. Children in this condition usually claimed that E1 had used a label consistent with a hybrid’s appearance rather than the one she had actually used just moments before. These errors are understandable in the context of the considerable information-processing demands of our task. First, by design, the hybrids did not look like what the experimenter claimed they were. For example, the key-like object had some features of a spoon, but it really did look like a key (as the Preliminary Study showed). Thus, children had to remember that an

object that looked very much like a key was actually a spoon, a task made even more challenging because the experimenter referred to it as a “spoon” just twice. Second, the experimenter referred to the typical items with labels that did match their appearance, so children had to keep track of which items in a particular set had been referred to as expected and which ones had not. Finally, children may have been distracted by the fact that the experimenter left the room after the inference task and before the naming one—a procedural decision we made in order to keep the memory control as close to the experimental condition as possible.

Other studies in which children learn something that is counter-intuitive have also shown that children sometimes later “misremember” this information (e.g., Kaiser, Proffitt, & McCloskey, 1985; Gopnik & Astington, 1988; Karmiloff-Smith, 1992; Moses & Flavell, 1990). For example, Gopnik and Sobel (2000, Study 1) found that 3- and 4-year-old children who watched as one novel object activated a “blicket detector” and a second, perceptually identical object did not, sometimes later mistakenly claimed that both had activated the detector. Gopnik and Sobel suggested that children in their study made the reasonable assumption that perceptually identical objects had similar causal powers—even though they had just seen that in the blicket detector case, they did not. They argue that “young children will reinterpret and even misremember events that took place only a moment or so before in the light of their higher level beliefs and expectations” (p. 1221).

In the present study, children in the memory control presumably had the reasonable expectation that an object’s appearance is a good cue as to what it is (e.g., Bloom, 2000), and this expectation over-rode the brief exposure children had to the unexpected label. A mechanistic account might involve something like the following<sup>2</sup>: Hearing E1 refer to the hybrid with an unexpected label primed a weak representation of the named category. This representation was active during the inference task, and children either relied on it or on the competing perceptually based representation, depending on how confident they were in their own judgment about the item’s category membership (and possibly in how credible they viewed E1 as being; see Jaswal & Malone, 2007). By the time the naming task took place, the effect of the priming had largely dissipated, often leaving children with the more salient perceptually based representation.

This does raise the question, of course, of what would be needed for children to retain the unexpected labels, or indeed, to retain other kinds of counter-intuitive information. The study described here was not designed to address this question, but we expect that if children had been exposed to the unexpected labels more frequently, they might have better been able to remember them. Additionally, studies in science education with older children suggest that explicitly noting the discrepancy between the information being conveyed and the children’s expectations can facilitate learning (e.g., Chi, 1992).

The study reported here is important because it validates a crucial assumption that researchers who use the category induction procedure to study the nature of children’s categories have made. When children make a label-based inference about stimuli that are moderately discrepant from their expectations, they seem to do so because they actually believe that label, not just because they are complying with the considerable experimental demands of the task.

## Acknowledgements

We thank the children, parents, and teachers who participated in these studies. Thanks also to Jennifer Bailey, Heather Burns, Lindsay Goldman, and Rachel Sylla for assistance in collecting and coding data, and to Dave McKercher for helpful comments on a previous draft. The research was supported by NICHD Grant HD-053403 to VKJ.

---

<sup>2</sup>We thank an anonymous reviewer for suggesting this possibility.

## References

- Akhtar N, Carpenter M, Tomasello M. The role of discourse novelty in early word learning. *Child Development* 1996;67:635–645.
- Aronsson K, Hundeide K. Relational rationality and children's interview responses. *Human Development* 2002;45:174–186.
- Bloom, P. How children learn the meanings of words. Cambridge, MA: MIT Press; 2000.
- Ceci SJ, Bruck M. Suggestibility of the child witness: A historical review and synthesis. *Psychological Bulletin* 1993;113:403–439. [PubMed: 8316609]
- Chi, MTH. Conceptual change within and across ontological categories: Examples from learning and discovery in science. In: Giere, RN., editor. *Minnesota studies in the philosophy of science, Volume 15: Cognitive models of science*. Minneapolis: University of Minnesota Press; 1992. p. 129-186.
- Clark, EV. The young word maker: A case study of innovation in the child's lexicon. In: Wanner, E.; Gleitman, LR., editors. *Language acquisition: The state of the art*. Cambridge, UK: Cambridge University Press; 1982.
- Clark, EV. Acquisitional principles in lexical development. In: Gelman, SA.; Byrnes, JP., editors. *Perspectives on language and thought: Interrelations in development*. Cambridge, UK: Cambridge University Press; 1991. p. 31-71.
- Csibra, G.; Gergely, G. Processes of change in brain and cognitive development. In: Munakata, Y.; Johnson, MH., editors. *Attention and performance XXI*. Oxford: Oxford University Press; 2006. p. 249-274.
- Dawkins R. What use is religion? Part 1. *Free Inquiry* 2004;24(4):13–14.
- Deutsch M, Gerard HB. A study of normative and informational social influences upon individual judgment. *Journal of Abnormal & Social Psychology* 1955;51:629–636.
- Diesendruck G, Markson L. Children's avoidance of lexical overlap: A pragmatic account. *Developmental Psychology* 2001;37:630–641. [PubMed: 11552759]
- Donaldson, M. *Children's minds*. New York: Norton; 1978.
- Gelman, SA. *The essential child: Origins of essentialism in everyday thought*. Oxford University Press: London; 2003.
- Gelman SA, Coley JD. The importance of knowing a dodo is a bird: Categories and inferences in 2-year-old children. *Developmental Psychology* 1990;26:796–804.
- Gelman SA, Markman EM. Categories and induction in young children. *Cognition* 1986;23:183–209. [PubMed: 3791915]
- Gelman SA, Markman EM. Young children's inductions from natural kinds: The role of categories and appearances. *Child Development* 1987;58:1532–1541. [PubMed: 3691200]
- Gelman SA, Waxman SR. Looking beyond looks: Comments on Sloutsky, Kloos, and Fisher (2007). *Psychological Science* 2007;18:554–555. [PubMed: 17576270]
- Gopnik A, Astington JW. Children's understanding of representational change and its relation to the understanding of false belief and the appearance-reality distinction. *Child Development* 1988;59:26–37. [PubMed: 3342716]
- Gopnik A, Sobel DM. Detectingblickets: How young children use information about novel causal powers in categorization and induction. *Child Development* 2000;71:1205–1222. [PubMed: 11108092]
- Graham SA, Kilbreath CS, Welder AN. Thirteen-month-olds rely on shared labels and shape similarity for inductive inferences. *Child Development* 2004;75:409–427. [PubMed: 15056196]
- Grice, HP. Logic and conversation. In: Cole, P.; Morgan, JL., editors. *Syntax and semantics, Volume 3: Speech acts*. New York: Academic Press; 1975. p. 41-58.
- Harris, PL. What do children learn from testimony?. In: Carruthers, P.; Stich, S.; Siegal, M., editors. *The cognitive basis of science*. Cambridge, UK: Cambridge University Press; 2002. p. 316-334.
- Henderson AME, Graham SA. Two-year-olds' appreciation of the shared nature of novel object labels. *Journal of Cognition and Development* 2005;6:381–402.
- Jaswal VK. Don't believe everything you hear: Preschoolers' sensitivity to speaker intent in category induction. *Child Development* 2004;75:1871–1885. [PubMed: 15566385]

- Jaswal VK. Preschoolers favor the creator's label when reasoning about an artifact's function. *Cognition* 2006;99:B83–B92. [PubMed: 16185679]
- Jaswal VK. The effect of vocabulary size on toddlers' receptiveness to unexpected testimony about category membership. *Infancy* 2007;12:169–187.
- Jaswal VK, Malone LS. Turning believers into skeptics: 3-year-olds' sensitivity to cues to speaker credibility. *Journal of Cognition and Development* 2007;8:263–283.
- Jaswal VK, Markman EM. Looks aren't everything: 24-month-olds' willingness to accept unexpected labels. *Journal of Cognition and Development* 2007;8:93–111.
- Jaswal VK, Neely LA. Adults don't always know best: Preschoolers use past reliability over age when learning new words. *Psychological Science* 2006;17:757–758. [PubMed: 16984291]
- Kaiser MK, Proffitt DR, McCloskey M. The development of beliefs about falling objects. *Perception and Psychophysics* 1985;38:533–539. [PubMed: 3834398]
- Karmiloff-Smith, A. *Beyond modularity: A developmental perspective on cognitive science*. Cambridge, MA: MIT Press; 1992.
- Kemler Nelson DC, Egan LC, Holt L. When children ask, 'what is it?' what do they want to know about artifacts? *Psychological Science* 2004;15:384–389. [PubMed: 15147491]
- Koenig M, Clement F, Harris PL. Trust in testimony: Children's use of true and false statements. *Psychological Science* 2004;15:694–698. [PubMed: 15447641]
- Koenig MA, Echols CH. Infants' understanding of false labeling events: The referential roles of words and the speakers who use them. *Cognition* 2003;87:179–208. [PubMed: 12684199]
- MacDonald G, Nail PR, Levy DA. Expanding the scope of the social response context model. *Basic and Applied Social Psychology* 2004;26:77–92.
- Mervis, CB.; Pani, JR.; Pani, AM. Transaction of child cognitive-linguistic abilities and adult input in the acquisition of lexical categories at the basic and subordinate levels. In: Rakison, DH.; Oakes, LM., editors. *Early category and concept development: Making sense of the blooming, buzzing confusion*. New York: Oxford University Press; 2003. p. 242-274.
- Moses LJ, Flavell JH. Inferring false beliefs from actions and reactions. *Child Development* 1990;61:929–945. [PubMed: 2209197]
- Mugny G. Compliance, conversion and the Asch paradigm. *European Journal of Social Psychology* 1984;14:353–368.
- Nail PR, MacDonald G, Levy DA. Proposal of a four-dimensional model of social response. *Psychological Bulletin* 2000;126:454–470. [PubMed: 10825785]
- Nelson, K. *Monographs of the Society for Research in Child Development*. 38. 1973. Structure and strategy in learning to talk; p. 1-2. Serial No. 149
- O'Neill D. Two-year-old children's sensitivity to a parent's knowledge state when making requests. *Child Development* 1996;67:659–677.
- Pea RD. Origins of verbal logic: Spontaneous denials by two- and three-year-olds. *Journal of Child Language* 1982;9:597–626. [PubMed: 7174759]
- Reid, T. *An inquiry into the human mind on the principles of common sense*. University Park: Pennsylvania State University; 1764/1997.
- Siegal, M. *Knowing children: Experiments in conversation and cognition*. 2. East Sussex, UK: Psychology Press; 1997.
- Siegal M, Surian L. Conceptual development and conversational understanding. *Trends in Cognitive Sciences* 2004;8:534–538. [PubMed: 15556022]
- Siegal M, Waters LJ, Dinwiddy LS. Misleading children: Causal attributions for inconsistency under repeated questioning. *Journal of Experimental Child Psychology* 1988;45:438–456.
- Sloutsky VM, Fisher AV. Induction and categorization in young children: A similarity-based model. *Journal of Experimental Psychology* 2004;133:166–188. [PubMed: 15149249]
- Sloutsky VM, Kloos H, Fisher AV. When looks are everything: Appearance similarity versus kind information in early induction. *Psychological Science* 2007a;18:179–185. [PubMed: 17425540]
- Sloutsky VM, Kloos H, Fisher AV. What's beyond looks? Reply to Gelman and Waxman. *Psychological Science* 2007b;18:556–557.

Tomasello M, Haberl K. Understanding attention: 12- and 18-month-olds know what is new for other persons. *Developmental Psychology* 2003;39:906–912. [PubMed: 12952402]



Key-like artifact

Label: "Spoon"



Car-like artifact

"Shoe"



Hat-like artifact

"Cup"

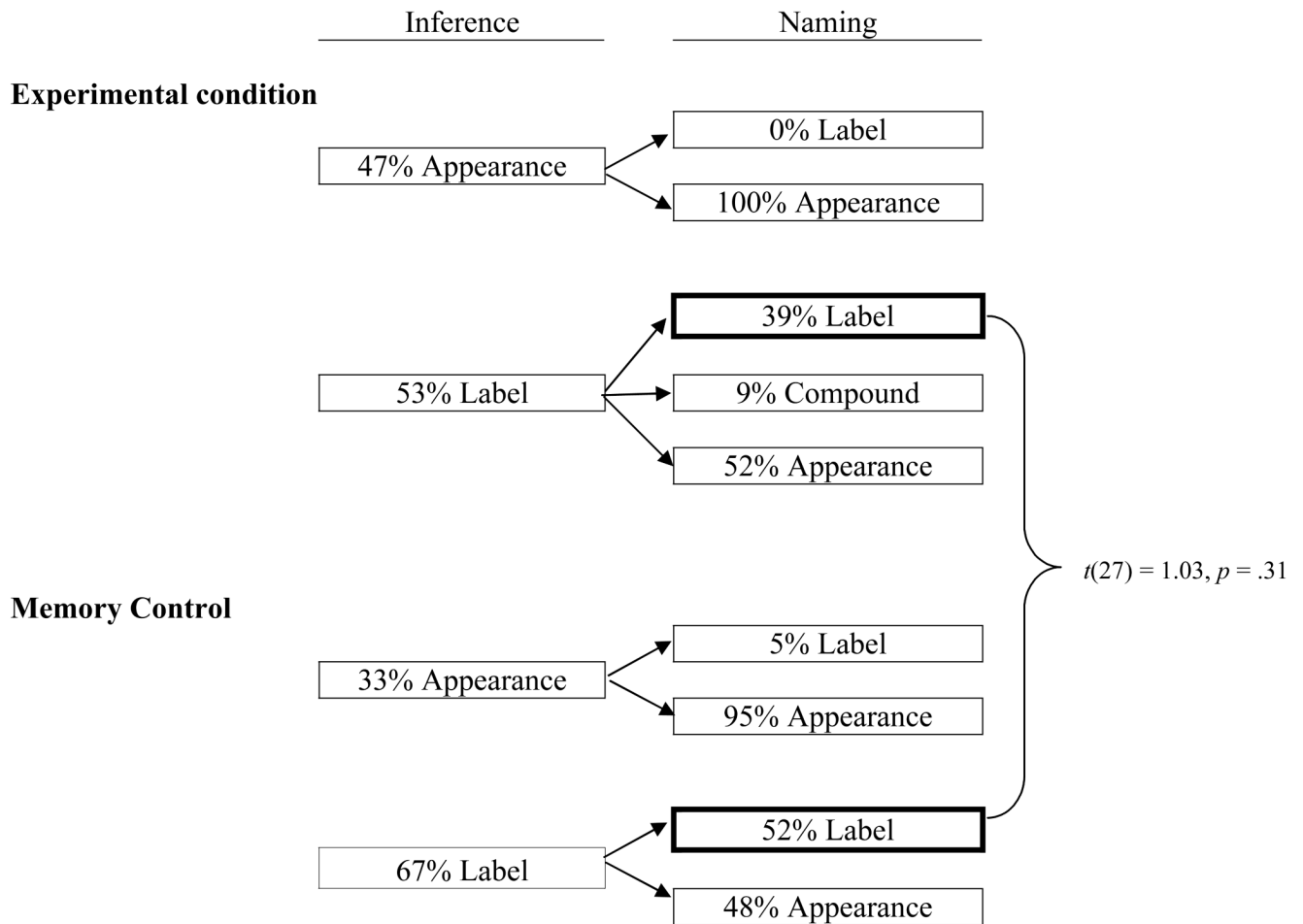


Toothbrush-like artifact

"Pen"

**Figure 1.**  
Hybrid artifacts. Children saw the stimuli in color.





**Figure 2.** Distribution of responses to the hybrid stimuli on the inference and naming tasks in each condition of the Primary Study. E1 asked the inference questions in both the Experimental and Memory Control conditions. E2 asked the naming questions in the Experimental condition; E1 did so in the Memory Control.

**Table 1**

## Stimulus Sets

Stimulus set	Function
Key Spoon	Starts the <b>car</b> <sup>a</sup> Eats the <b>cereal</b>
Hat Cup	Goes on the <b>man's head</b> Goes on the <b>table</b>
Toothbrush Pen	Brushes the <b>lady's teeth</b> Writes on the <b>paper</b>
Car Shoe	Goes on the <b>street</b> Goes on the <b>baby's foot</b>

<sup>a</sup>Bold indicates the background photograph used during the inference task.

**Table 2**  
 Number of Children Reporting 0, 1, 2, or 3 of the 4 Unexpected Labels to E2 (Experimental Condition) or E1 (Memory Control)

	# initial label-based inferences	# unexpected labels reported			
		0	1	2	3
Experimental Condition	0	2			
	1	3			
	2	1	3	1	
	3	0	1	2	
	4	0	2	0	1
Memory Control Condition	0	1			
	1	1			
	2	2	1	2	
	3	0	1	2	
	4	0	2	2	2