DEVELOPMENT OF SPATIAL REASONING IN YOUNG CHILDREN

by

SARAH ELIZABETH CUMMINS-SEBREE

(Under the Direction of Dorothy M. Fragaszy)

ABSTRACT

Spatial reasoning is the ability to know how objects are distributed and can move through space. Researchers who study this development have disagreed on when these skills are evident, possibly due to differing task presentations and behavioral requirements in these studies. This study was designed to test which of two theories provides a better explanation of the development of spatial reasoning in young children: the core knowledge hypothesis and an ecological approach. I presented 24-, 34-, and 44-month-old children a single task that required different behavioral responses based on a three-process (*recognition*, *production*, and *prediction*) spatial reasoning model grounded in Gibsonian ecological theory. In this task, I presented a box containing 1-3 vertical and diagonal translucent tubes leading to opaque containers to the children. Children were required to give three different behavioral responses: a) find a dropped ball after it had fallen through a tube (*recognition*), b) put a ball into the tube that led to a selected container (*production*), and c) indicate the end-point of a ball before it was dropped into a tube (*prediction*). Older children chose correctly more often than younger children for all tasks. No main effect for task presentation order was found (i.e., being able to produce the actions did not lead to better predictive performance). Multiple interactions indicated that differences in performance across tasks were due in part to age, gender, task presentation order, and the number of tubes presented. In addition, children who used a finger to trace the path of a tube in at least one trial during testing made more correct prediction choices than those who never traced tube paths. These findings suggest that producing strategic actions during these tasks facilitates the development of spatial reasoning in young children.

INDEX WORDS: Spatial reasoning, Child development, Gibsonian ecological theory, Affordances, Object knowledge

DEVELOPMENT OF SPATIAL REASONING IN YOUNG CHILDREN

by

SARAH ELIZABETH CUMMINS-SEBREE

B.A., Thomas More College, 1995

M.S., The University of Georgia, 1999

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2003

DEVELOPMENT OF SPATIAL REASONING IN YOUNG CHILDREN

by

SARAH ELIZABETH CUMMINS-SEBREE

Major Professor:	Dorothy M. Fragaszy
Committee:	James Brown Janet Frick Chester Karwoski Katherine Kipp

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
May 2003

DEDICATION

I dedicate this dissertation to my loving and VERY patient family. This includes my parents, Kathleen Cummins and Steve and Deborah Cummins, as well as my brother, Nathan Cummins. Their support and cheerleading allowed me to keep going through this long and arduous process.

But most of all, this work is dedicated to my incredibly supportive husband, Jerry Sebree. His unending confidence in me, and his willingness to stick it out, even when I was in tears, made this journey's completion possible. I love you, Jer!

ACKNOWLEDGEMENTS

I want to take this opportunity to thank Dr. Dorothy M. Fragaszy for her guidance and tutelage through my graduate career (I finally made it!). I also want to thank Drs. James Brown, Janet Frick, Chester Karwoski, and Katherine Kipp for being kind enough to serve on my doctoral committee and their helpful comments on my dissertation.

I want to thank Ron Davis at the University of Georgia Instrument Shop for constructing my apparatus, as well as other pieces of equipment I have needed during my graduate career. I also would like to acknowledge those who have assisted me in data collection and scoring: Stephanie Bock, Monique Dase, Katie Jackson, Meena Lambha, Carrie Rosengart, Tara Rountree, and Leslie Taylor.

Thanks must also go to the children who participated, as well as their parents who granted permission for their children to take part in this study. Many thanks also go to the daycare centers that allowed me to do my testing there: McPhaul Children's Center at the University of Georgia, Oconee Preschool Academy, Inc., Trinity Lutheran Daycare Center, and Young World Learning Center.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE

INTRODUCTION

Spatial reasoning is the cognitive ability to know how objects are distributed and how they can move through space (Tomasello & Call, 1997). How humans develop their spatial reasoning abilities has been widely studied in child research over the past fifty years. During that time, disagreement has arisen concerning what age particular reasoning skills (e.g. the ability to predict the end-point of a moving object) are evident (Hood, 1995, 1998, 2001; Hood, Carey, & Prasada, 2000; Spelke, Breinlinger, Macomber, & Jacobson, 1992). This disagreement is in part due to differing task presentations and behavioral requirements in these studies. Comparative studies on spatial reasoning with nonhuman primates have yet to resolve these issues (Hauser, Williams, Kralik, & Moskovitz, 2001; Hood, Hauser, Anderson, & Santos, 1999; Santos & Hauser, 2002), though I believe they may hold clues as to how to approach spatial reasoning development in young children (Cummins-Sebree & Fragaszy, 2000). The purpose of this study was to investigate the development of spatial reasoning skills in children aged 24 to 44 months by presenting tasks that require different behavioral responses. The study was designed to test hypotheses derived from a tri-part processing model of the development of spatial reasoning elaborated below. The proposition of three differing spatial reasoning processes in the model (recognition, production, and prediction) results from a synthesis of previous studies on spatial reasoning with human and nonhuman primates and their seemingly contradictory findings.

Previous Research

Piaget (1954) proposed that children construct their knowledge of objects and how these objects can move in the environment by interacting with those objects. During the fourth and fifth stages of the sensorimotor period (8-18 months), infants manipulate objects singly or with other objects to discover object properties and how different actions on objects (e.g. dropping) produce respective events (e.g. hitting the floor). At the completion of the sensorimotor period at about 2 years, the child has the ability to represent objects and what can be done with those objects mentally, and the child can predict object movement and end-point locations.

Subsequent research with infants has questioned the timing of the development of these abilities. Spelke and colleagues (1992) have developed the core knowledge hypothesis: infants are born with, or develop in the first few months of life, a set of core beliefs and the ability to reason about objects in space. This hypothesis is based on studies utilizing the preferential-looking paradigm. It rests on findings that show that infants older than 3-4 months will look longer at visual stimuli that are unusual or new to them. Infants look longer at impossible events (e.g. a falling ball appears below a solid platform as its trajectory end-point) than at possible events (e.g. a falling object appears on top of a solid platform) (Baillargeon, 1994; Kim & Spelke, 1992; Spelke et al., 1992). The infants in the previous studies range in age from 3.5 to 8 months. The investigators interpret longer looking times as indicating “surprise” at the impossible events. Spelke and colleagues (1992) conclude that these infants can predict end-points of moving objects based on object properties and spatial reasoning (and thus are surprised by violation of expectations) at an earlier age than Piaget suggested.

Hood (1995), however, has found that young children (some up to 48 months) exhibit difficulty in predicting end-points of moving objects in a search task wherein they are dropped into opaque tubes. Children ages 24 to 48 months exhibit gravity errors (i.e. they anticipate that all objects fall in a straight vertical fashion regardless of direction of supports (Hood, 1995), and 24-month-old children also disregard barriers as impediments to object movement (Hood, Carey, & Prasada, 2000). When directly compared on looking time and search tasks, 30- and 36-month-old children look longer at impossible events, but only the 36-month-olds search correctly (Hood, Cole-Davies, & Dias, 2003). Hood and colleagues (2000, 2003) conclude that infants and young children have not developed representations of object properties and their possible movements in space to enable accurate predictions of end-point locations. These findings seem to contradict those of Spelke and colleagues (1992); how can younger infants “predict” end-point locations of moving objects, but 2- and some 3- and 4-year-olds cannot?

Several differences between these two testing paradigms probably contribute to the conflicting findings. First, the tasks used to examine spatial reasoning have used different visual presentations and required different behavioral responses. The use of preferential-looking requires only visual attention as a response, and the presentation of an event allows viewing the progression of the moving object until it comes within a small distance from the support or barrier. At that time, an occluder appears, allowing the experimenter to place the object in the possible or impossible location (but see Hood et al., 2003 for a different presentation format). In the tube tasks, the object cannot be viewed as it falls, and the child is required to point to or tell the experimenter the expected end-point. Second, the preferential-looking paradigm used by Spelke et al.

(1992) required habituation with possible events before presenting probe and control trials to determine preference; habituation to different trial types can affect infant visual preferences (e.g. Kannass and Oakes, 1997; Schilling, 2000). Few familiarization trials are used in the tube prediction tasks. As a result of these differences in task presentation and behavioral responses, different spatial reasoning processes may be measured. A model of spatial reasoning that specifies recognition, production, and prediction as distinctive abilities permits a tentative resolution to the apparently discordant results from these studies.

Three-Process Model of Spatial Reasoning: Recognition, Production, and Prediction

The three-process model of spatial reasoning is a synthesis of the abilities of young children evident in the work reviewed above. Conceptually, this model is based on the ecological theory of perceiving affordances in the environment and its relation to one's developing action repertoire (Gibson & Pick, 2000). The model conceptualizes spatial reasoning as encompassing these partially dependent processes: recognition, prediction, and production. These hypothetical processes are defined and operationalized procedurally as follows:

Recognition is defined as the ability to attend to the physical features of events within view and differentiate between them and other events that have been seen. This only requires an individual to detect the physical differences between two events and respond accordingly. When presented with an event to watch that differs significantly from previously seen events, an individual who recognizes this differences would presumably look longer at that event, regardless of knowledge of physical laws. This would be an accurate explanation for the results obtained in preferential-looking tasks, such as those by Spelke et al. (1992).

Prediction is defined as inferring what can happen next in an event sequence when one is not controlling that event sequence. This requires an individual to be capable of detecting the affordances in the environment that allow certain events to occur and that preclude other events from occurring, and also to infer from the current affordances what will happen next. The finding that younger children have difficulty in predicting the end-point location of a ball traveling through an opaque tube (Hood, 1995) or behind solid doors (Hood et al., 2003) can be explained as a failure to detect the affordances of the tubes or barriers between doors and/or to infer the final location of the ball.

Neither of the commonly used paradigms taps a third process: production. *Production* is defined as using a motor behavior to solve a spatial problem. This requires an individual to detect the affordances of the immediate environment and employ the correct motor action to attain a goal. This process may be the ability under development as describe in Piaget's (1954) observations of his children repeatedly dropping and manipulating objects in different ways to see what the outcomes would be; for example, his children would often attempt to retrieve an out-of-reach object with an inappropriate tool before learning that a different object is more appropriate (e.g. using a blanket to retrieve a ball under the sofa before figuring out that a broom works better).

Studies with capuchin monkeys suggest that these three processes can be dissociated (Cummins-Sebree & Fragaszy, 2000). Capuchins presented with modified versions of Spelke et al.'s tasks (1992) chose correctly a potential end-point location of a circling ball when a hole or barrier was a feature of the supporting platform (recognition).

However, they chose incorrectly the end-point location of a moving ball that had reached a hole in the supporting platform (prediction). These capuchins also learned over testing to avoid barriers and holes (attending to barriers earlier than holes) when using a hoe to retrieve a food treat (production) (Cummins, 1999).

Ecological Theory: The Basis of the Three-Process Model

Each of these processes involves varying abilities to detect affordances of the present situation and to respond to moving objects accordingly (Gibson, 1979). Detecting affordances requires perceiving the structures of objects and surfaces in relation to the goals and action capabilities of the individual (Gibson, 1988; Gibson & Pick, 2000). Each process requires perceiving particular information and exhibiting a distinctive behavioral response. For recognition, one must only detect physical differences to respond with “surprise” at an event that looks different from what has been seen in the past. For production, one must detect the physical features of the situation and produce appropriate motor actions to solve a problem. To be capable of prediction, one must perceive the physical features of the situation, infer the next event in an event sequence, and indicate that to another individual, either through vocalization or motoric action (e.g. pointing, searching).

E. J. Gibson and Pick (2000) outlined a theory of perceptual development based on the ecological theory of perception developed by J. J. Gibson (1979). One’s ability to detect affordances in the environment is contingent upon the individual’s ability to act on that environment. For example, crawlers and experienced walkers perceive different surfaces as traversable (Adolph, 1997). When crawlers begin walking, they attempt to do so on surfaces they have previously crawled on successfully. However, they discover quickly that a slope that allows crawling does not necessarily allow walking (i.e. it is too steep). Walkers learn to detect new affordances (what makes a traversable surface based on their walking expertise) only after implementing their newly developed action pattern.

I believe that this is also the case for the spatial reasoning processes outlined above. Based on Spelke et al. (1992) and Baillargeon's (1994) work, I hypothesize that we develop our recognition ability early in life. For recognition, we only need to be able to follow visually a moving object and remember the event we have seen. If a different event is presented and we are able to detect the differences in the presentation, then we will look longer at the novel event. This action pattern (looking longer at a novel event) is well developed in the first few months of life (Colombo, 1993; Gibson & Pick, 2000). As we develop physically over our first year, our abilities to move our bodies and, consequently, to detect more information in our environment increase. Thus, we develop our ability for production; we do this through active exploration of the objects in our environment and how they behave based on actions done on them (i.e. Piaget, 1954). Through our increasing ability to produce appropriate responses to solve spatial problems and to learn from those actions, we become able to infer the endings of event sequences (thus, capable of prediction), such as the end-point location of an unseen moving ball based on the contours of its container (i.e. Hood, 1995), and we can do this without being an agent in the event sequence. It is through the development of our action repertoire that we are able to learn to detect new affordances in the environment and apply what we have learned in more complex contexts in the future (Gibson & Pick, 2000). (For further review of the literature, please see Appendix A.)

Purpose, Methodology, and Hypotheses

The purpose of my study was to a) evaluate the existence of these distinctive spatial reasoning processes in young children and b) identify their trajectory of development. By investigating the existence of recognition, production, and prediction processes and their developmental trajectory, one can determine if the core knowledge hypothesis or the ecological approach to perceptual development better represents the contradictory findings in the spatial reasoning literature. I used a variation of Hood's (1995) task to do this, and I selected children aged 24-, 34-, and 44-months to participate as these age groups represent the age range that have had varied levels of difficulty in his studies. I presented the preschoolers with an apparatus consisting of up to three pink, translucent, vertically diagonal tubes that led to up to three gray opaque containers. Each child was asked to provide three different responses in this task: a) to watch a blue ball (dropped by me) fall through a tube and then indicate through pointing or search the container in which the ball fell (recognition- Rec), b) to put the ball into a tube so that it would fall into a designated container (production-Pro), and c) to point to or open the container door corresponding to the end of the tube from which I was holding the ball at the top end of the tube before dropping it (prediction-Pre). I also included a set of transfer (Tra) trials in the prediction format to investigate the children's abilities to transfer their knowledge to novel tube configurations. The dependent variables were first choice of container (recognition, prediction, transfer) or tube (production) and the number of attempts required for the children to correctly place the ball into the designated tube (production).

By requiring three different behavioral responses, many hypotheses could be generated to test which theory (core knowledge hypothesis or ecological theory) better accounts for spatial development in young children. If the core knowledge hypothesis is the better theoretical model, then all children would perform equally well on the prediction and recognition tasks, regardless of age or number of tubes presented; this is because those that propose the core knowledge hypothesis believe that the preferential-looking measures indicate predictions on the correct end-point locations of moving objects. Though the core knowledge hypothesis only provides an expectation concerning the recognition and prediction abilities, all children should be successful at the production tasks under this hypothesis, as even 24-month-olds have the motoric ability to place smaller objects into larger ones.

If the ecological theory is the better theoretical model, then there should be differences in prediction abilities compared to recognition abilities not only based on age, but also by the order in which the children are required to make the different responses and the number of tubes presented. Younger children should perform better on the recognition task compared to the prediction task, whereas older children should perform equally well on both tasks. However, these differences should be tempered by the order in which they receive the tasks; those children who are asked to produce the actions before predicting end-points (Rec-Pro-Pre: Order 1) should perform better on the prediction task than those who are asked to predict the end-points before producing the actions (Rec-Pre-Pro: Order 2) or are never asked to produce the actions across testing (Rec-Rec-Pre: Order 3). Children who are not asked for the production response at all should perform more poorly on the transfer tasks than either of the other two groups of children who are asked for the production response at some point during testing. Also, younger children should require more attempts at producing the correct responses in the production task than older children. Increasing the number of tubes may also increase the visual complexity of the arrangement, making it more difficult to detect the affordances of the correct tube; Hood (1995) found that increasing the number of tubes from one to three in the apparatus yielded increased numbers of errors. (For an illustration of these hypotheses, see Table 1.1.)

One dependent variable and two independent variables were tested in this study in addition to those given above. First, Hood (1995, 1998) found that the errors children made during search tasks reflected a gravity bias – the idea that all objects fall straight down. When a child made this error, he searched in the location directly below the top of a diagonal tube from which the ball was dropped, not in the correct location. Thus, I hypothesized that the errors made by all children would consist of these gravity errors. Second, some spatial reasoning studies (e.g. Hood et al., 2003) have found gender differences in spatial reasoning skills, with males outperforming females. Thus, I included gender in my analyses. Third, a participant early in the testing period exhibited a behavior that may have helped him make correct choices during the prediction and transfer phases. This behavior, which I named tracing, consisted of using a finger to trace the outline of the tube from outside of the box before making a choice; this tracing could occur for any portion of the tube. Thus, I hypothesized that those children who spontaneously traced the pattern of a tube at least one time during testing would make more correct choices during prediction and transfer trials than those children who did not spontaneously trace.

Table 1.1

Hypotheses Based on Competing Theories

Theory		Task		
	Recognition	Production	Prediction	Transfer
Core Knowledge Hypothesis	24 = 34 = 44 in first choice	24 = 34 = 44 in first choice and attempts	24 = 34 = 44 in first choice	24 = 34 = 44 in first choice
	Order 1 = Order 2 = Order 3 in first choices		Order 1 = Order 2 = Order 3 in first choices	Order 1 = Order 2 = Order 3 in first choices
	Recognition = Prediction = Transfer in first choices		Recognition = Prediction = Transfer in first choices	Recognition = Prediction = Transfer in first choices
Ecological Theory	24 = 34 = 44 in first choice	24 < 34 = 44 in first choice and attempts	24 < 34 < 44 in first choice	24 < 34 < 44 in first choice
	Order 1 = Order 2 = Order 3 in first choices	Order 1 = Order 2 in first choices and attempts	Order 1 > Order 2 = Order 3 in first choices	Order 1 = Order 2 > Order 3 in first choices

CHAPTER TWO

METHOD

In this study, I tested the spatial reasoning skills of 24-, 34-, and 44-month-old children by using a modified version of Hood's (1995) tube task. Each participant was asked to exhibit three different behavioral responses during testing: *recognizing* the end-point of a ball that has traveled through a tube, *producing* the correct action in putting a ball into a tube that leads to an open door, and *predicting* which container a ball will travel to when dropped into a tube.

Participants

Forty-three participants were recruited from four daycare centers in the Athens, GA area. Information letters and permission slips were sent home with the children in the appropriate age range, and those children whose parents sent back the permission slips participated in testing. Three participants' data were not included in the analyses due to their inability or lack of cooperation to complete at least half of the experimental tasks; all three were 24-month-olds. Thus, the final data set (40 participants) consisted of 11 24-month-olds (M = 24.15, SD = .40; range: 23.67 – 24.97; nine males, two females), 14 34-month-olds (M = 33.92, SD = .35; range: 33.53 – 34.57; three males, 11 females), and 15 44-month-olds (M = 43.96, SD = .40; range: 43.47 – 44.90; seven males, eight females). These participants consisted of 19 males and 21 females; 30 were Caucasian, 5 were Hispanic, 3 were Asian, and 2 were African-American.

Materials

A clear PVC box with an open back (5-sided box) contained three holes at the top and three opaque containers with doors at the bottom of the box. The doors opened outward toward the participants so that they could search for the fallen ball within the containers. Three pink translucent tubes directed the paths that a blue clay ball could take from the top of the box to each of the opaque containers (see Figure 2.1). A second clear PVC box, similar in design but containing only one tube and one container, was used to familiarize the child with the task for that testing session.

Figure 2.1. Photograph of participant working on the transfer trials.

An RCA Pro846 camcorder and 8 mm videotapes were used to record all trials for scoring of the data. Participants received a sticker at the beginning and end of each testing day, and they also received an age-appropriate toy and certificate of completion at the end of the final day of testing.

Procedure

Before testing, each child was asked to watch a blue ball roll through a horizontal tube three to four times while I watched the child for visual tracking of the moving ball; this was to ensure the child was capable of the visual requirements needed for the tasks (all children passed this task). Also, each child received two to four familiarization trials with the training apparatus (single tube and single container) to ensure the child would understand what behavior was required for that day's testing.

For each task, testing trials were presented with one to three tubes present (Hood, 1995). Five trials for each number of tubes were administered, each with a different configuration of tubes, for a total of 15 trials. Participants received the one-tube trials first, followed by two and then three tubes for each task (except for the transfer set). The tube configurations were presented randomly within the tube number set; for the 15 trials, four consisted of vertical tubes used as the correct tubes (two for the one tube set and one for the two and three tube sets), and 11 consisted of diagonal tubes as the correct tubes. For the transfer set, all five trials consisted of diagonal tubes as the correct tubes.

Participants were assigned to one of three task order groups in a counterbalanced fashion across age groups. Those in the recognition-production-prediction (Order 1) group reflected the ecological theory hypothesis, and those in the recognition-prediction-production (Order 2) group reflected the core knowledge hypothesis; the third group (recognition-recognition-prediction: Order 3) became the control group to determine if receiving the recognition task twice would equal the effects of producing the actions (as could be seen in the Order 1 group) before the prediction task. Each task required 10-25 minutes of testing (depending on the child's level of interest and ability to choose the correct tube or container), and testing occurred over a three-day period.

Recognition Task. The child was seated at a table or on the floor, facing the apparatus. I opened the doors to the containers to show the child that they could open for retrieval of the ball. After getting the child's attention, I dropped the ball into the selected tube while the child watched. The child was then asked to find the ball. If the child did not find the ball on the first try, she was allowed to continue searching until the ball was found. First choice of container, the correctness of that choice, occurrence of the gravity error, and any incidence of tracing was recorded.

Production Task. The procedure was similar to the one noted in the previous task. However, I opened a container door, handed the ball to the child, and then asked the child to put the ball at the top so that it would fall down into the container with the opened door. If the child placed the ball into the correct tube, he was given verbal praise and we continued to the next trial. If the child placed the ball into the incorrect tube, the child was asked to try again; this continued until the child placed the ball into the correct tube. First choice of container, the correctness of that choice, occurrence of the gravity error, the number of attempts required needed to succeed on each trial, and any incidence of tracing was recorded.

Prediction Task. The procedure was similar to the one noted in the first task. However, I held the ball at the top of a tube and asked the child to open the door to the container that the ball would go to when I dropped it. After the child made his decision (correct or incorrect), I dropped the ball and asked the child to retrieve it. First choice of container, the correctness of that choice, occurrence of the gravity error, and any incidence of tracing was recorded.

Transfer Trials. The procedure was the same as that of the prediction task, but five novel configurations were used. Again, first choice of container, the correctness of that choice, occurrence of the gravity error, and any incidence of tracing was recorded.

Scoring. Data were scored from videotape or live during testing by an undergraduate assistant. Interrater reliability ranged from 86% to 100% on correctness of choice on all three tasks and the transfer trials and number of attempts required during the production task.

CHAPTER 3

RESULTS

I compared the proportion of correct first choices for the three tasks and the transfer trials across age, gender, task, task order, and number of tubes present. Average number of attempts required for the production task was compared across age, gender, and task order. Gravity errors were compared across age, gender, and task order. I subdivided participants into groups reflecting the occurrence of tracing on at least one trial during testing (tracers and non-tracers), and differences between the two groups were compared on their proportion of correct first choices for the prediction and transfer tasks. All analyses were performed using SPSS 11.0 software; I set alpha at $p = .05$. Fisher's LSD was used for post hoc analyses involving repeated measures variables, and Tukey's HSD was used for post hoc analyses involving main effects for between-subjects variables.

Recognition and Prediction

The core knowledge hypothesis leads to the prediction that the proportion of correct first choices should not differ between recognition and prediction. The ecological theory, however, predicts that there should be age, task order, and task effects on the proportion of correct first choices. All children should perform equally on the recognition task, whereas older children should perform better on the prediction task. Those in the Order 1 group should perform better on the prediction task than the other two groups.

Performance should decline with increasing numbers of tubes, due to the visual complexity of the tubes (i.e. requiring extra attention to the affordances of the tubes).

A 3 x 2 x 3 x 2 x 3 (Age x Gender x Order x Task x Number of Tubes) Mixed Analysis of Variance (ANOVA) was performed to determine the effects of those variables on the proportion of correct first choices. For all significant effects, power values ranged from .62 to 1.00. There were main effects for Age, Task, and Number of Tubes (see Table 3.1). Post hoc tests indicated that 24-month-olds (M = .50, SD = .13) made significantly fewer correct choices 44-month-olds (M = .86, SD = .13), and the difference between the 24-month-olds and 34-month-olds (M = .76, SEM = .15) approached significance (see Table 3.2). Significantly more correct first choices were made in the recognition task (M = .78, SD = .14) than in the prediction task (M = .69, SD = .19) (see Table 3.1). Significantly fewer correct first choices were made for trials with three tubes (M = .68, SD = .21) compared to one tube (M = .78, SD = .17) and two tubes (M = .74, SD = .14) (see Table 3.2).

However, these main effects are tempered by the multiple significant interactions that occurred. There were a number of two-way interactions that were significant. First, there was a significant Task x Age interaction (F (2, 19) = 6.34, p = .008). Seven of 11 24-month-olds performed above chance levels on the recognition task (with three possible choices, chance is p = .33), whereas all 34- and 44-month-olds performed above chance. In the prediction task, nine of 11 24-month-olds, 12 of 14 34-month-olds, and all 44-month-olds performed above chance levels. As seen in Figure 3.1, the trends suggested that 24-month-olds performed more poorly than the 34- and 44-month-olds on the recognition and prediction tasks; the 34-month-olds performed equally compared to

Table 3.1

ANOVA for Correct First Responses for Recognition and Prediction

Source	*df*	*MSE*	F
Age	2	1.77	17.96
Task	1	.72	19.93
Task x Age	2	.23	6.34
Task x Gender	1	.64	17.73
Task x Age x Gender	2	.13	3.56
Task x Gender x Order	2	.30	8.29
Tubes	2	.18	6.57
Task x Tubes	2	.20	10.10
Task x Tubes x Age	4	.12	5.90
Task x Tubes x Age x Gender	4	.09	4.29

Note. All *F* values presented are significant to $p = .05$.

Table 3.2

Post Hoc Tests of Main Effects for Recognition and Prediction Analyses

Variable	Comparison	df	*t*	*p*
Age	24 mo. and 44 mo.	20	6.08	<.001
	24 mo. and 34 mo.	18	4.36	.004
Tubes	1 Tube and 3 Tubes	35	3.23	.003
	2 Tubes and 3 Tubes	35	2.86	.007

Note. All *p* values are significant (less than .05).

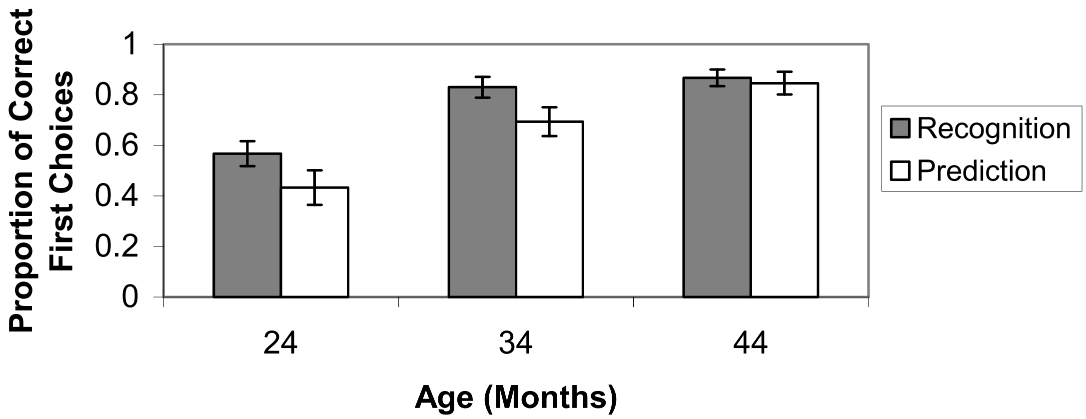


Figure 3.1. Proportion of correct first choices on the recognition and prediction tasks across ages (24 months, 34 months, 44 months).

the 44-month-olds on the recognition task, but they performed more poorly compared to the 44-month-olds on the prediction task. Post hoc tests indicated that the 24-month-olds performed significantly more poorly on both tasks compared to the 44-month-olds, as well as compared to the 34-month-olds on the recognition task. The comparisons between the 24- and 34-month-olds and the 34- and 44-month-olds were also significant (see Table 3.3).

Second, there was a significant Task x Gender interaction (F (1, 19) = 17.73, $p <$ .001). The trends suggested that males performed equally on the recognition and prediction tasks; however, females performed better on the recognition task than on the prediction task, as well as outperformed the males on the recognition task (see Figure 3.2). Post hoc tests indicated that females performed significantly better on the recognition task compared to the prediction task ($t(19) = 4.51$, $p < .001$); no other comparisons approached significance.

Table 3.3

Post Hoc Tests for the Task x Age Interaction for Recognition and Prediction Analyses

Task	Comparison	df	*t*	*p*
Recognition	24 mo. and 44 mo.	20	4.53	<.001
	24 mo. and 34 mo.	18	4.29	<.001
Prediction	24 mo. and 44 mo.	20	4.74	<.001
	24 mo. and 34 mo.	18	2.16	.039
	34 mo. and 44 mo.	26	3.07	.004

Note. All *p* values are less than .05.

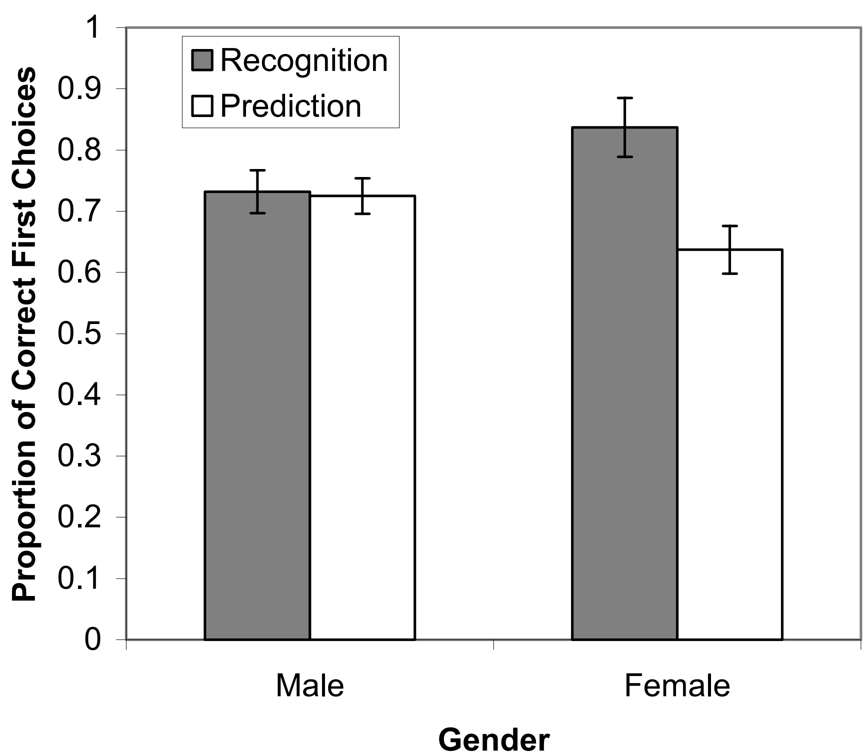


Figure 3.2. Proportion of correct first responses for the recognition and prediction tasks for each gender.

Third, a significant Task x Number of Tubes interaction occurred (F (2, 38) = 10.10, $p < .001$). According to Figure 3.3, the trends suggested that the number of tubes presented did not affect performance in the recognition task; however, performance declined as the number of tubes went from one to two in the prediction task. Post hoc tests indicated significant differences between one-tube and three-tube performance for the prediction task, and significant differences between the two tasks for the two-tube presentations and the three-tube presentations (see Table 3.4).

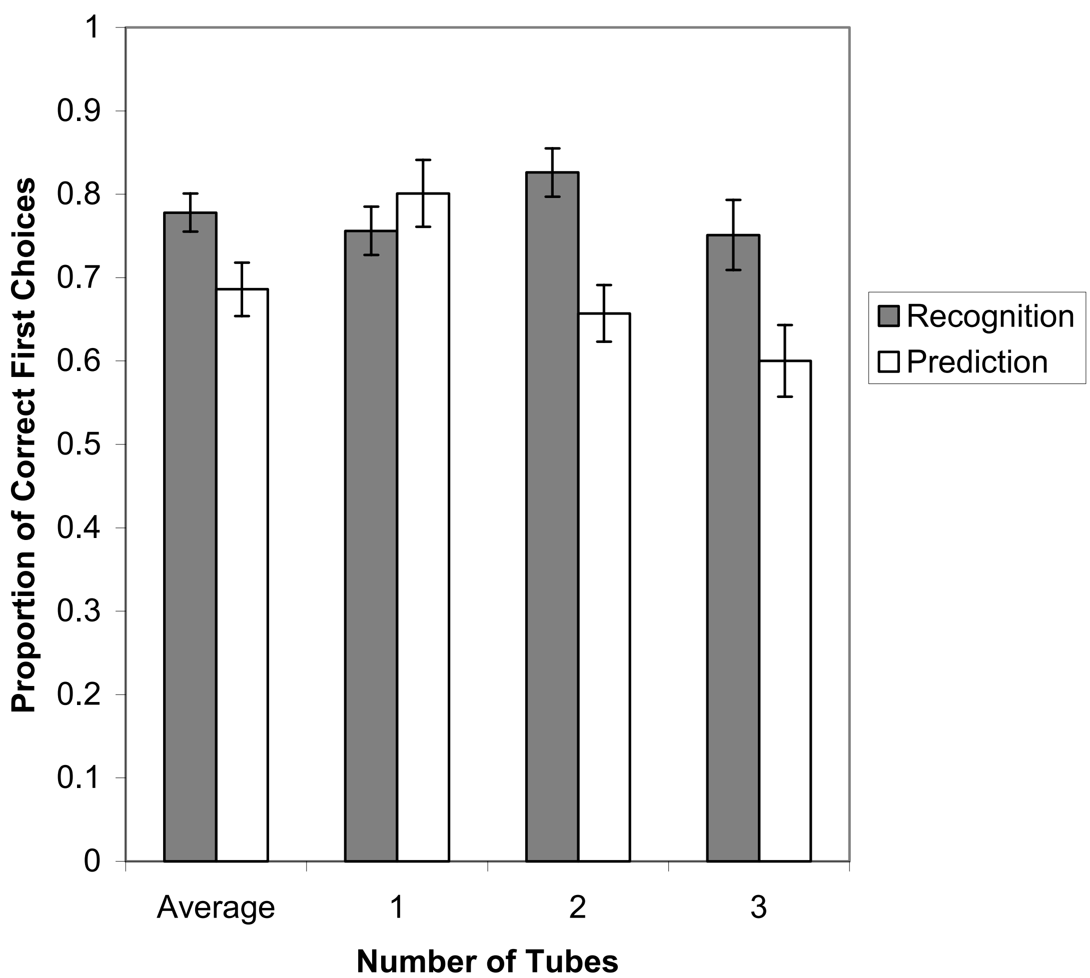


Figure 3.3. Proportion of correct first choices for the recognition and prediction tasks across the number of tubes presented.

Table 3.4

Post Hoc Tests for the Task x Tube Interaction for Recognition and Prediction Analyses

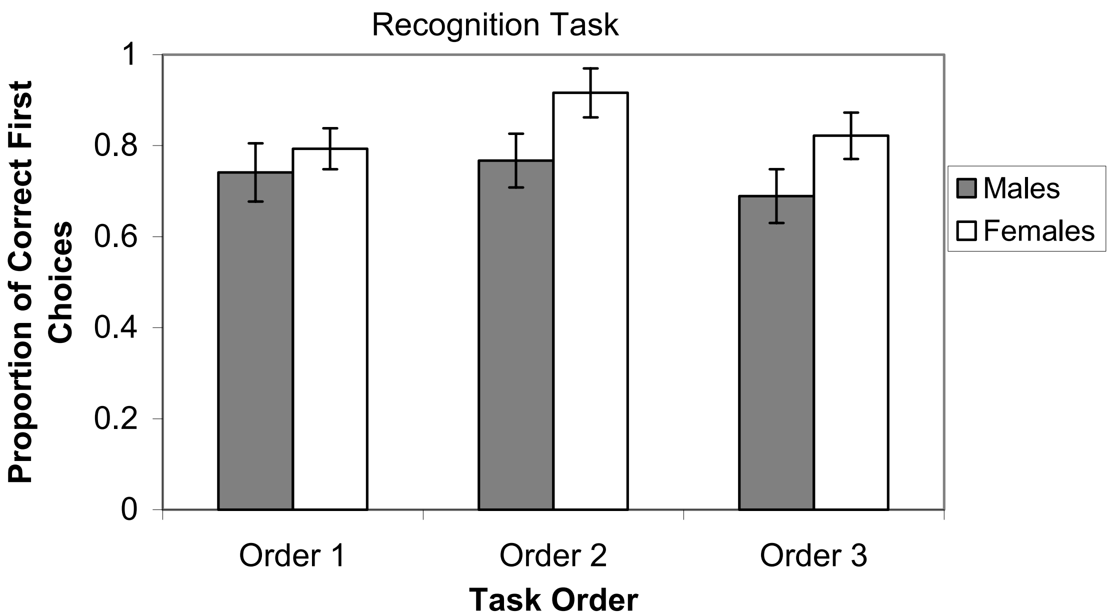
Task	Comparison	df	*t*	*p*
Prediction	1 Tube and 3 Tubes	34	4.97	<.001
	1 Tube and 2 Tubes	34	3.24	.003
Both	2 Tubes	34	3.38	.002
	3 Tubes	34	3.66	.001

Note. All *p* values are less than .05.

Additionally, there were significant three-way interactions. First, there was a significant Task x Gender x Order interaction (F (2, 19) = 8.29, p = .003). As seen in Figure 3.4 (a, b), the trends suggested that females in the Order 2 group outperformed the males in all order groups on the recognition task, whereas the males in the Order 2 group outperformed the females in all order groups on the prediction task. Post hoc analyses indicated that the difference in performance on the recognition task between males and females in Order 2 was significant (t(9) = 2.35, p = .025). There were also significant differences in performance on the recognition and prediction tasks for females Order 1 (t(7) = 2.84, p = .007), in Order 2 (t(5) = 2.90, p = .007), and in Order 3 (t(5) = 3.15, p = .003). Though the females in all three presentation orders performed significantly worse on the prediction task than the recognition task, the decrease in performance was less than half for females in Order 1 (11%) than those in Order 2 (26%) or Order 3 (27%).

Second, there was a significant Task x Gender x Age interaction (F (2, 19) = 3.56, p = .05). As seen in Figure 3.5 (a, b), the trends suggested that 24- month-olds males and

A)



B)

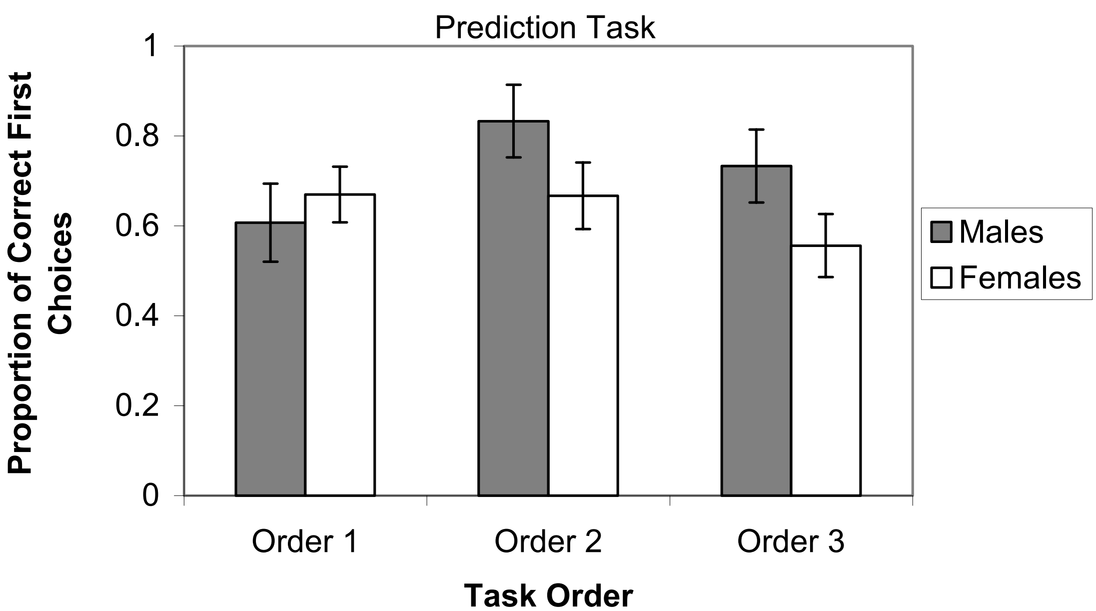


Figure 3.4 a, b. Proportion of correct first choices for recognition (a) and prediction (b) as a function of gender and task order. Descriptions of task order can be found in the methods section.

A)

B)

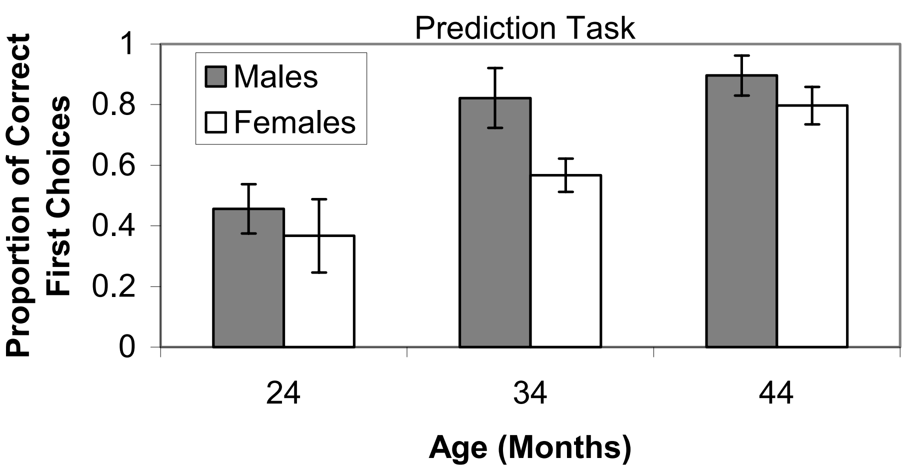


Figure 3.5 a, b. Proportion of correct first responses for the recognition (a) and prediction (b) tasks as a function of age (24, 34, 44 months) and gender.

females performed more poorly on both the recognition and prediction tasks compared to the 34- and 44-month-olds; however, 34-month-olds females performed more poorly on the prediction tasks than on the recognition tasks, whereas 34-month-olds males and 44-

month-olds of both genders did equally well on the recognition and prediction tasks. Post hoc analyses indicated that 34-month-old females performed significantly worse on the prediction task than on the recognition task ($t(9) = 4.93$, $p < .001$).

Third, there was a significant Task x Age x Number of Tubes interaction (F (4, 38) = 5.90, $p = .001$). As seen in Figure 3.6 (a, b), the trends suggested that the 24-month-olds exhibit variable performance across one, two, and three tubes in the recognition task, and they show a decrease in correct answers in the prediction task when moving from one to two tubes. Thirty-four-month-olds exhibit a gradual decline in performance over one, two, and three tubes in the prediction task, but not in the recognition task; 44-month-olds exhibit comparable performance across tubes and task. Post hoc analyses for this interaction indicated that both 24- and 34-month-olds performed significantly worse on the prediction task compared to the recognition task when two tubes were present, and 34-month-olds performed significantly worse on the two tasks when three tubes were present (see Table 3.5). All other comparisons were contained within other interactions.

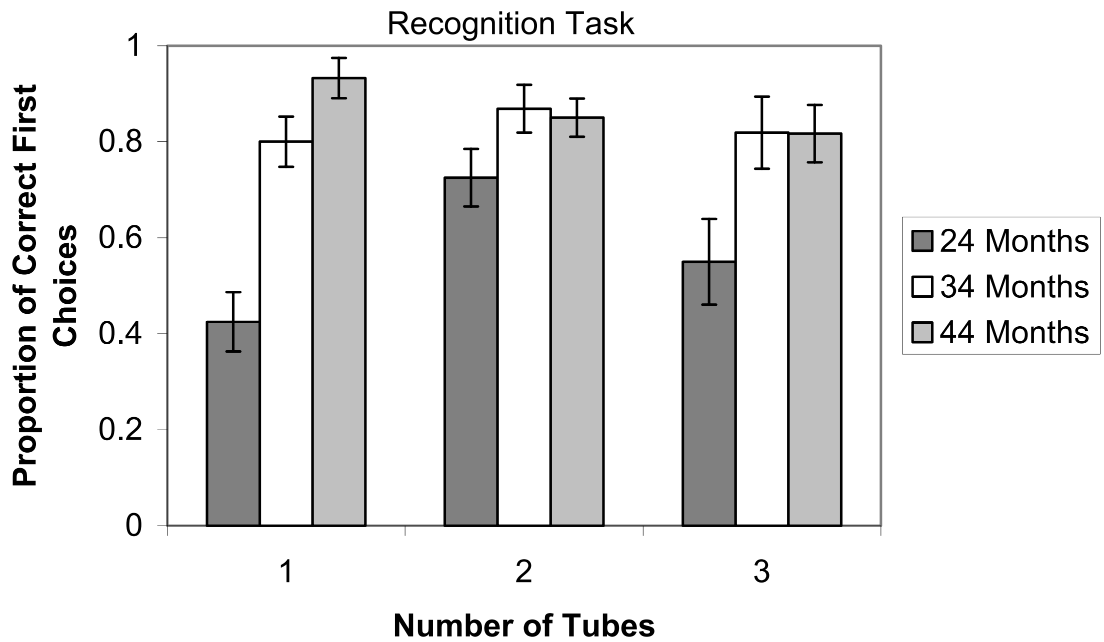
Table 3.5

Post Hoc Tests for the Task x Age x Tube Interaction for Recognition and Prediction Analyses

Age	Comparison	df	*t*	*p*
24 months	2 Tubes	6	4.16	<.001
34 months	2 Tubes	12	4.05	<.001
	3 Tubes	12	4.72	<.001

Note. Alpha was set to $p = .05$ for these analyses.

A)



B)

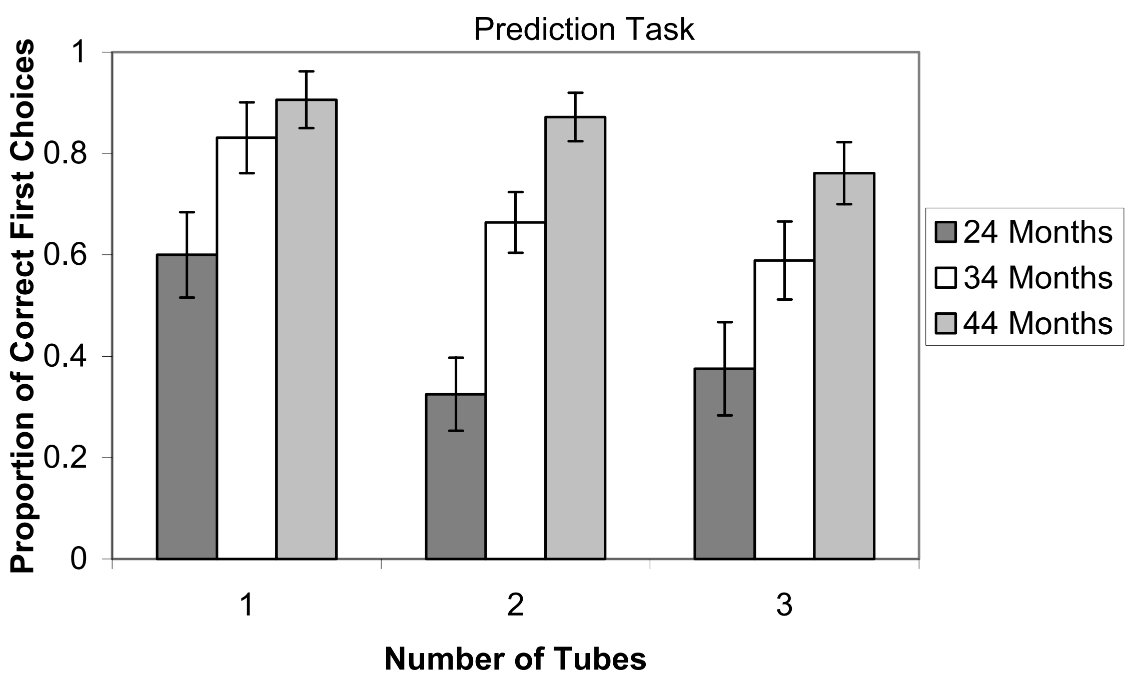


Figure 3.6 a, b. Proportion of correct responses for recognition (a) and prediction (b) tasks as a function of age (24, 34, 44 months) and the number of tubes presented.

There was also a four-way interaction for Task x Tubes x Age x Gender (see Table 3.1). As these types of interactions rarely illuminate any additional information above and beyond multiple three-way interactions, I will not discuss this interaction.

Production

The core knowledge hypothesis posits that there should be no differences due to age or task order for the proportion of correct first choices or average number of attempts required to succeed at the production task. The ecological theory, however, predicts that younger children would differ in the proportion of correct first responses made and the average number of attempts needed for success compared to older children; it also predicts that increasing the number of tubes would affect those dependent variables by increasing the visual complexity of the apparatus layout.

Correct Choices. A 3 x 2 x 2 x 3 (Age x Gender x Order x Number of Tubes) ANOVA was performed to determine the effects of those variables on the proportion of correct first choices for the production task. (The production choices were separated from the other two tasks due to the automatic removal of the Order 3 group during analyses by SPSS; because of this and the requirement for multiple post hoc comparisons, alpha was set to $p = .005$). Power values for all significant effects ranged from .63 to 1.00. There were main effects for Age and Number of Tubes (see Table 3.6). Twenty-four-month-olds ($M = .62$, $SD = .13$) made significantly fewer correct choices than 44-month-olds ($M = .87$, $SD = .13$), and this was true for the difference in performance between 34-month-olds ($M = .64$, $SD = .15$) and 44-month-olds (see Table 3.7). Six of seven 24-month-olds performed above chance levels, and all 34- and 44-month-olds made correct choices above chance levels. More correct first choices were made with one tube ($M = .94$, $SD =$

Table 3.6

ANOVA for Proportion of Correct First Responses for Production

Source	*df*	*MSE*	*F*
Age	2	.44	10.29
Tubes	2	.97	22.96
Tubes x Gender x Order	2	.17	3.99
Tubes x Age x Gender x Order	2	.16	3.67

Note. All *F* values presented are significant to $p = .05$.

Table 3.7

Post Hoc Tests for the Main Effects for Production Analyses

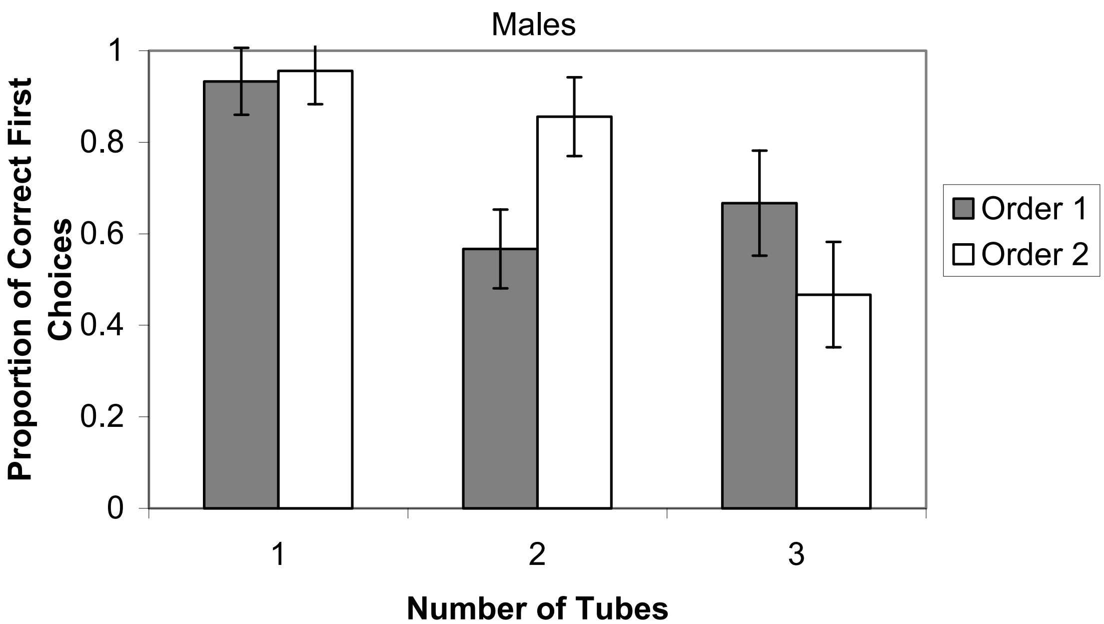
Variable	Comparison	df	*t*	*p*
Age	24 mo. and 44 mo.	15	4.59	.001
	34 mo. and 44 mo.	17	3.88	.004
Tubes	1 Tube and 2 Tubes	25	5.21	<.001
	1 Tube and 3 Tubes	25	10.42	<.001
	2 Tubes and 3 Tubes	25	5.21	.004

Note. Alpha was set to $p = .05$ for these analyses.

.18) than with two tubes (*M* = .72, *SD* = .21); both one and two tubes yielded more correct responses than three tubes (*M* = .50, *SD* = .28) (see Table 3.7).

Additionally, there was a significant Gender x Order x Number of Tubes interaction (F (2, 30) = 3.99, p = .03). As seen in Figure 3.7 (a, b), the trends suggest that

A)



B)

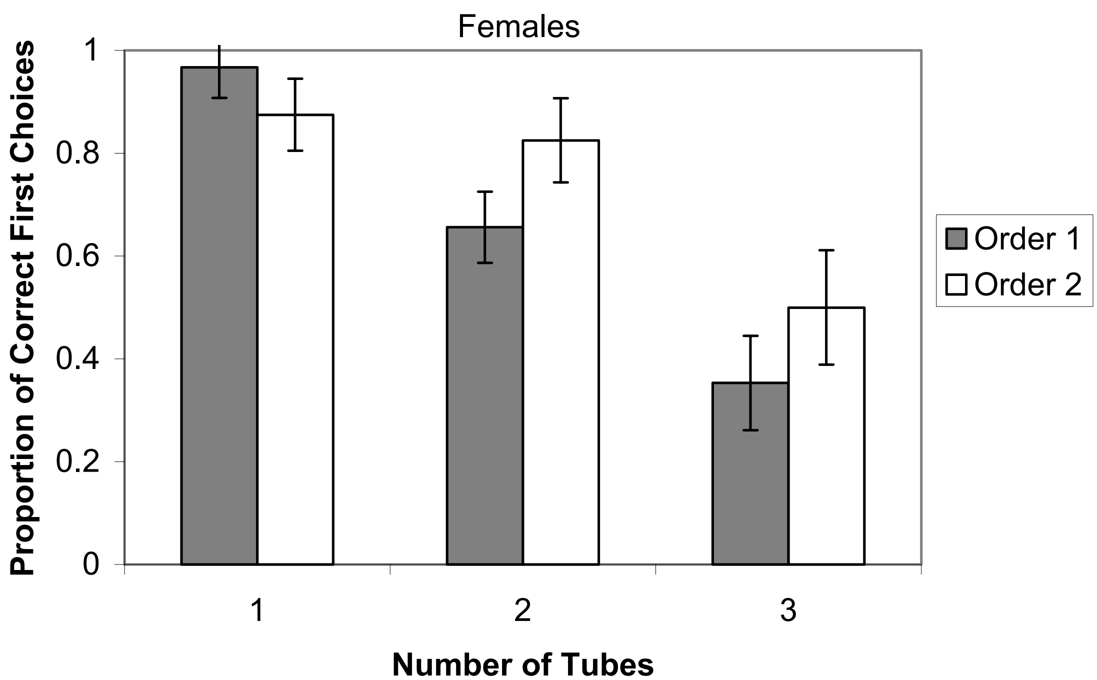


Figure 3.7 a, b. Proportion of correct first choices for the production task as a function of task order and number of tubes for males (a) and females (b).

males and females in the Order 2 group make comparable correct first choices for one and two tube presentations and made fewer correct first choices for the three tube presentations. It also appeared that males in the Order 1 group made comparable correct first choices with two- and three-tube configurations, both of which were fewer than those for one tube; females of that group exhibited decreases in correct first choices when moving from one tube to two tubes to three tubes. Post hoc tests indicated that both genders in Order 1 made significantly more correct choices when one tube was presented than when two tubes were presented ($t(5) = 5.13$, $p < .001$ for males; $t(7) = 5.12$, $p < .001$ for females), but only the difference in performance between two tubes and three tubes was significant for females in that order ($t(7) = 2.50$, $p = .02$). In Order 2, males made significantly fewer correct choices with 1-tube presentations than with 3-tube presentations ($t(5) = 2.26$, $p = .03$); this also occurred for the females in this order ($t(5) = 3.88$, $p = .001$), but they also performed significantly worse with 2-tube presentations compared to 3-tube presentations ($t(5) = 3.37$, $p = .003$).

There was also a significant Tubes x Age x Gender x Order interaction. As these types of interactions rarely illuminate any additional information above and beyond multiple three-way interactions, I will not discuss this interaction.

Average Attempts Required. There were no significant main effects for Age, Gender, or Order, nor were there any significant interactions when comparing those variables on the average number of attempts required for success. However, the ecological theory posits that younger children should require more attempts to succeed at production trials than older children. Therefore, a planned comparison was done, examining any potential differences between the 24-month-olds and the 34- and 44-

month-olds. As expected, 24-month-olds (*M* = 2.36, *SD* = 1.21) required more attempts to be successful compared to the 44-month-olds (*M* = 1.19, *SD* = .13) (F (2, 24) = 6.82, p = .005).

Transfer Trials

As these trials are simply prediction trials that are novel to the participants, the predictions given for the core knowledge hypothesis for the prediction task would be identical here. Under the ecological theory, those in the Order 1 and Order 2 conditions should make more correct first choices on the transfer trials than those in the Order 3 condition, due to their ability to produce the actions of dropping the ball at some point during testing; it would also predict that older children would make more correct first responses than younger children due to increased experience with working with objects during the participants' lifetimes.

A 3 x 2 x 3 (Age x Gender x Order) ANOVA was conducted to investigate the effects of those variables on the transfer prediction trials (see Table 3.8). There was a significant main effect for Age; 44-month-olds (*M* = .82, *SD* = .27) made significantly more correct first choices than 24-month-olds (*M* = .24, *SD* = .28) and 34-month-olds (*M* = .58, *SD* = .32) ($t(21) = 4.91$, $p < .001$ for 24-month-olds; $t(19) = 3.24$, $p = .01$ for 34-month-olds). Two of eight 24-month-olds, ten of 13 34-month-olds, and 13 of 15 44-month-olds made correct choices above chance levels. There were no significant interactions for this set of trials.

Gravity Errors

As Hood (1995, 1998) found that errors made by young children in spatial reasoning tasks tend to be gravity errors (i.e. selecting the location directly under the

Table 3.8

ANOVA for Proportion of Correct First Responses for Transfer

Source	*df*	*MSE*	*F*
Age	2	1.06	12.71*
Gender	1	.17	.98
Order	2	.03	.88
Age x Gender	2	.06	.75
Age x Order	4	.08	1.10
Gender x Order	2	.19	2.48
Age x Gender x Order	2	.08	1.02

Note. Those *F* values marked with an asterisk are significant to $p = .05$.

point of launching or dropping the object), it would be expected that these types of errors would occur in this study as well. A 3 x 2 x 3 (Age x Gender x Order) ANOVA was conducted using the average number of gravity errors made across testing as the dependent variable. No main effects or interactions were significant (see Table 3.9). Interestingly, group means ranged from .32 (males in the Order 2 group) to 1.00 (34-month-old males in the Order 3 group). Most of the means ranged from .40 to .60. The expected frequency of gravity errors would be .50 (one container is the correct choice, with the two other containers consisting of the gravity error and non-gravity error).

Tracing

As mentioned previously, some participants used a finger to trace the figure of the tube from the outside of the apparatus during testing. Those who exhibited this behavior

Table 3.9

ANOVA for Proportion of Gravity Errors Made

Source	*df*	*MSE*	*F*
Age	2	.03	.50
Gender	1	.01	.07
Order	2	.08	.32
Age x Gender	2	.03	.96
Age x Order	4	.05	1.72
Gender x Order	2	.16	4.60
Age x Gender x Order	2	.03	.56

Note. No *F* values were significant at $p = .05$.

at least once were labeled as tracers, and those who did not were called non-tracers. The core knowledge hypothesis makes no prediction concerning this behavior; however, the ecological theory posits that this behavior may assist the child in extracting the affordance of the tube, and thus increase the number of correct first choices compared to those children who did not trace the outline of the tube.

Sixteen (ten female, six male) of the 40 children exhibited the tracing behavior at some point during testing. Two of the tracers were 24-months-olds, five were 34-months-old, and nine were 44-months-old. Independent t-tests were conducted with tracer (yes or no) as the independent variable and the overall proportion of correct first choices for recognition, production, prediction, transfer trials, and gravity errors as the dependent variables. Only two the prediction and transfer variables yielded significant results, and the recognition variable approached significance. Tracers made more correct first choices than non-tracers on the prediction task and on the transfer trials, and this trend also occurred with the recognition task (see Table 3.10).

Table 3.10

Differences Between Tracers and Non-tracers on Proportion of Correct First Choices and Gravity Errors

Dependent Measure	Tracers (*M, SD*)	Non-tracers (*M, SD*)	*df*	*t*
Recognition	.82, .04	.70, .05	38	1.84
Production	.69, .06	.69, .04	25	.001
Prediction	.75, .06	.59, .04	38	2.30*
Transfer	.74, .09	.45, .07	34	25.9*
Gravity Errors	.54, .06	.53, .05	38	.12

Note. The *t* values marked with an asterisk are significant at $p = .05$.

CHAPTER 4

DISCUSSION

Evidence for the Ecological Theory (Three-Process Model)

The bulk of the results supported the ecological theory. First, age affected the proportion of correct first choices across the all tasks. Twenty-four-month-olds made more errors on the recognition and prediction tasks than the 34- and 44-month-olds; 34-month-olds performed equally well compared to the 44-month-olds on the recognition task, but not on the prediction task. The two younger age groups performed worse than the oldest group on the production task. Additionally, all three age groups differed on the proportion of correct responses for the transfer trials (novel prediction trials), in which there is a gradual increase in correct choices across age.

In fact, the patterns of the proportion of correct responses differed for the three age groups. Twenty-four-month-olds were similar in their modest recognition (57%) and production (62%) choices, but exhibited a decline in performance through the prediction (43%) and transfer (24%) tasks. Thirty-four-month-olds exhibited a high performance rate for the recognition task (83%), but made roughly 60% correct choices for the other three tasks (production: 64%; prediction: 69%; transfer: 58%). Forty-four-month-olds made correct choices at similar proportions across all tasks (recognition: 87%; production: 87%; prediction: 85%; transfer: 82%). Thus, recognition abilities seemed mature or become stable at age 34 months, and production and prediction abilities continue to develop up past that age, but with different rates.

Second, the proportion of correct first choices and average number of attempts required to succeed on the production task differed across age. Twenty-four- and 34-month-olds made fewer correct first choices than 44-month-olds, and 24-month-olds required more attempts to succeed than 44-month-olds. Therefore, production abilities seemed to increase steadily over this 20-month period. However, we can assume that this ability is still under development, as even the 44-month-olds continued to make a few errors.

Third, the number of tubes affected the number of correct first choices for the prediction and production tasks; as the number of tubes increased, the number of correct first choices decreased. This effect was greater for the 24-month-olds than for the 34-month-olds, whereas the 44-month-olds exhibited equivalent performance across tasks and tube numbers.

The predicted task order effect did not occur. Overall, children who received the production task prior to the prediction task did not make more correct choices on the prediction task compared to those who either received the production task later in testing or did not receive the production task at all. In addition, this pattern did not occur when considering the age groups separately. Those who received the production task did not perform better on the transfer trials than those who never received the production task; again, this pattern did not occur within age groups. Thus, it seemed that across or within age groups, producing the action of dropping the ball into the tubes did not facilitate prediction abilities (in the prediction or transfer tasks).

Evidence for the Core Knowledge Hypothesis

Very little evidence emerged to support the core knowledge hypothesis. The evidence that did appear involved the lack of task order effects for the recognition and prediction tasks and overall lack of effects for age on the number of attempts required to succeed in the production task. Thus, receiving the production task before the prediction task did not facilitate an overall increase in the proportion of correct first choices made across age groups, and 34-months-olds require fewer attempts than the 24-month-olds or more attempts than the 44-month-olds on the production task. There were also no task order effects for the transfer trials; having received the production task at some point in time did not provide an overall benefit when receiving the transfer trials compared to those children who did not receive the production task during testing. Additionally, the lack of differences in gravity errors and the fact that proportions of errors that were gravity errors across conditions (.40 to .60) were not different from chance (.50) supported the core knowledge hypothesis stance. According to this hypothesis, children this age have expectations of how objects move; thus, any errors made should be randomly distributed between two incorrect containers, and not based on a gravity rule that all objects fall in a vertical fashion.

Effects Due to Gender: The Interactions

Many significant interactions occurred with gender as one of the defining variables. Though a main effect for gender did not occur, the interactions were moderately strong (21-46% of the variance accounted for included gender as a variable). Males performed equally well on the recognition and prediction tasks, whereas females seemed to perform better than the males on the recognition task but worse than the males on the prediction task. This was further complicated by many factors. First, age influenced children's performances such that 24-month-old males and females performed equally poorly on both tasks compared to the older age groups, and 44-month-old males and females performed equally well on both tasks compared to the youngest age group. The 34-month-old females exhibited a decrease in correct first choices in the prediction task, and the 34-month-old males did not. Thus, there seemed to be a lag in the development of the prediction ability for the females across this age range that is not seen in males; males developed this ability by 34 months, while the females developed this ability by 44 months.

Second, these gender differences for the recognition and prediction tasks were also affected by the task order. Whereas males performed equally well across tasks regardless of the presentation order in which they were placed, females who were given the production task before the prediction task exhibited a proportionally smaller decline in correct first responses compared to the females who either were given the production task after the prediction task or were not given the production task at all during testing. Thus, it seemed that females were able to improve performance across tasks if they were allowed to put the ball in the tubes before being required to make predictions (as posited in the ecological theory).

Third, gender interacted with task order and the number of tubes presented during the production task. Males and females who were given the production task after the prediction task exhibited a decline in correct first choices when moving from two tubes to three tubes. Males who were given the production task before the prediction task gave fewer correct responses when moving from one to two tubes, but did not provide more incorrect responses when moving from two to three tubes. Females who were in the same group, however, exhibited an equal decline in correct first choices when moving from one to two to three tubes.

Tracing: Additional Evidence Supporting the Ecological Theory

As noted previously, some children spontaneously used a finger to trace the outline of the correct tube on the front of the apparatus before making a choice. Those who traced at least once during testing made more correct first choices during the prediction and transfer trials than those who did not trace. Thus, tracing may have assisted the children in detecting the affordances of the tubes so that they could make the correct choice. Though this behavior is not the production behavior hypothesized to yield increases in correct prediction responses, it is still a motoric behavior that children used as a strategy to generate the correct answer to a problem; therefore, the tracing behavior falls under the definition of a production ability (using a motor behavior to solve a spatial problem).

Core Knowledge Hypothesis vs. Ecological Theory

When comparing the evidence for the core knowledge hypothesis and the ecological theory (three-process model), the results support the ecological theory as a more complete explanation for the development of spatial reasoning in young children. The age and task effects found for the three different tasks provide evidence that recognition, production, and prediction are three distinct abilities. Thus, it is reasonable to propose that the studies by Spelke et al. (1992, 1994) and Baillargeon (1994) may be measuring recognition, whereas the studies by Hood and colleagues (1995, 1998, 2000) may be measuring prediction.

In addition, performances on these different tasks exhibit different developmental trajectories, which may be timed differently for males and females. Recognition develops before production and prediction, and it may be mature by around age 34 months for males and females. Prediction may develop much later for females (around 44 months) than it does for males (around 34 months). During this time, production may develop more gradually between 24 and 44 months for both genders.

Also, the ability to produce other behaviors may aid in the development of the prediction ability. Tracers spontaneously produced a motor behavior that allowed them to follow the shape and direction of tube in action. Perhaps this made it easier to pick up the visual information that specified the tube's affordances. As Gibson & Pick noted (2000), the development of our action capabilities enables us to detect new affordances in the environment and apply what we have learned in more complex contexts in the future. Tracers had developed an appropriate action that allowed them to detect the affordances of the tubes in the more complex prediction task.

Implications for Further Research and Conclusions

The findings presented here provide support for the notion that spatial reasoning development is more complicated than the core knowledge hypothesis would suggest. Regardless of the overall differences in task performance due to age, the multiple interactions indicate that developmental researchers should further investigate the role of gender, visual complexity in task presentation, and the requirements of differing behavioral responses in the development of spatial reasoning. In addition, more attention should be given towards discerning the behaviors that young children use to explore the spatial properties and relations among objects in the environment. Researchers should also investigate how these behaviors may influence spatial reasoning development through increasing a child's abilities to detect affordance information. More specifically, future research should focus on the development of the tracing behavior seen in this study. By requiring some participants to trace tubes and preventing others from doing so, we may be better able to understand the extent to which that behavior assists in spatial reasoning development or indexes a substantive change in spatial reasoning.

REFERENCES

Adolph, K. E. (1997). Learning in the development of infant locomotion. *Monographs of the Society for Research in Child Development, 62* (3, Serial No. 251).

Baillargeon, R. (1994). Physical reasoning in young infants: Seeking explanations for impossible events. *British Journal of Developmental Psychology, 12*, 9-33.

Cohen, L. B., Gilbert, K. M., & Brown, P. S. (1996, April). *Infants' understanding of solidity: Replicating a failure to replicat*e. Poster presented at International Conference on Infant Studies, Providence, RI.

Colombo, J. (1993). *Infant cognition: Predicting later intellectual functioning*. London, UK: SAGE Publications, Inc.

Cummins, S. E. (1999). Detection of environmental constraints in a tool-use task by tufted capuchin monkeys (*Cebus apella*). Unpublished master's thesis, University of Georgia, Athens.

Cummins-Sebree, S. E., & Fragaszy, D. M. (2000, April). *Can capuchin monkeys predict trajectories of moving objects?* Paper presented at Southern Society for Philosophy and Psychology conference, Atlanta, Georgia.

Gibson, E. J. (1988). Exploratory behavior in the development of perceiving, acting, and the acquiring of knowledge. *Annual Review of Psychology, 39*, 1-41.

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

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

Hauser, M. D., Williams, T., Kralik, J. D., & Moskovitz, D. (2001). What guides search for food that has disappeared? Experiments on cotton-top tamarins (*Saguinus oedipus*). *Journal of Comparative Psychology, 115*, 140-151.

Hood, B., Carey, S., & Prasada, S. (2000). Predicting the outcomes of physical events: Two-year-olds fail to reveal knowledge of solidity and support. *Child Development, 71*, 1540-1554.

Hood, B., Cole-Davies, V., & Dias, M. (2003). Looking and searching measures of object knowledge in preschool children. *Developmental Psychology, 39*, 61-70.

Hood, B. M. (1995). Gravity rules for 2- to 4-year-olds? *Cognitive Development, 10*, 577-598.

Hood, B. M. (1998). Gravity does rule for falling events. *Developmental Science, 1*, 59-63.

Hood, B. M. (2001). When do infants know about objects? *Perception, 30*, 1281-1284.

Hood, B. M., Hauser, M. D., Anderson, L., & Santos, L. (1999). Gravity biases in a non-human primate? *Developmental Science, 2*, 35-41.

Kannass, K. N., & Oakes, L. M. (1997, April). *Down hill vs. up hill: Infants' sensitivity to gravitational constraints on object motion*. Poster presented at The Biennial Meeting of the Society for Research in Child Development, Washington, D. C.

Kim, I. K., & Spelke, E. S. (1992). Infants' sensitivity to effects of gravity on visible object motion. *Journal of Experimental Psychology: Human Perception and Performance, 18*, 385-393.

Piaget, J. (1954). *The construction of reality in the child*. New York: Basic Books, Inc.

Santos, L. R., & Hauser, M. D. (2002). A non-human primate's understanding of solidity: dissociations between seeing and acting. *Developmental Science, 5*, F1-F7.

Schilling, T. H. (2000). Infants' looking at possible and impossible screen rotations: The role of familiarization. *Infancy, 1*, 389-402.

Spelke, E. S. (2000). Core knowledge. *American Psychologist, 55*, 1233-1243.

Spelke, E. S., Breinlinger, K., Macomber, J., & Jacobson, K. (1992). Origins of knowledge. *Psychological Review, 99*, 605-632.

Spelke, E. S., Katz, G., Purcell, S. E., Ehrlich, S. M., & Breinlinger, K. (1994). Early knowledge of object motion: continuity and inertia. *Cognition, 51*, 131-176.

Tomasello, M. & Call, J. (1997). *Primate cognition*. New York: Oxford University Press.

APPENDIX A

LITERATURE REVIEW

Spatial reasoning is the ability to understand how objects can be distributed in and move through space (Tomasello & Call, 1997). Organisms reason about space when they move competently in their environment and utilize and predict object movement (e.g. predators must predict the trajectory of prey to catch their next meal; a child must be able to understand that her sled will fall into a hole in the ice covering a pond if it continues toward that hole). Are we born with this ability, or does it develop over time? If it develops over time, when do we become skilled at spatial reasoning?

Sensorimotor Period: Constructing Spatial Knowledge in a Piagetian Model

Piaget proposed that young children develop their spatial reasoning skills by acting on objects in their environment. According to Piaget (1954), acting on objects is understood only in relation to the child as agent during Stage IV of the sensorimotor period (around ages 8-12 months). For example, if a child pushes a spoon off of a table, she recognizes the movement of the spoon as resulting from her action; she does not view it as the spoon is falling due to a lack of support from the table. As the child executes more actions on objects in more novel ways (e.g. dropping the spoon from different positions, pushing the spoon down inclines, etc.), she begins to understand the relations among objects outside of her own actions. This understanding represents a key accomplishment of Stage V of the sensorimotor period (ages 12-18 months).

Preferential-Looking Paradigm and the Core Knowledge Hypothesis

Piaget's (1954) observations with his children reflected their spatial reasoning abilities as determined by their motoric actions, such as reaching, throwing, and searching. Because younger infants are not physically capable of such actions, a new method of testing their spatial reasoning abilities was needed. The preferential-looking paradigm has become a popular method to investigate physical reasoning because it does not require pointing or other arm or hand movements, and vision is a primary sensory modality for young infants.

The preferential-looking paradigm is based on evidence from infant attention studies that indicate infants look longer at novel events than at familiar events. In infant spatial reasoning studies (e.g. Baillargeon, 1994; Spelke, Breinlinger, Macomber, & Jacobson, 1992; Spelke, Katz, Purcell, Ehrlich, & Breinlinger, 1994), infants are repeatedly shown an event until habituation occurs; following this, novel exemplars that reflect "possible" or "impossible" events are presented. If the novel "possible" events do not promote longer looking, whereas the "impossible" novel events do, the conclusion is drawn that infants are sensitive to (i.e. are "surprised" by) violations of physical laws (e.g. solidity, gravity).

Spelke and colleagues (1992) proposed the core knowledge hypothesis as a result of a collection of experiments with young infants that utilized the preferential-looking paradigm. The core knowledge hypothesis states that infants are born with, or develop very early in life, basic spatial reasoning skills related to object movement. These include reasoning about continuity (i.e. objects will travel in straight lines until acted on by another object or surface), solidity (i.e. a moving object cannot move through another solid object), and gravity (i.e. objects lacking support from a surface will fall). As we develop, we are able to reason about more complex object movement (e.g. planetary orbits, trajectories of billiard balls based on the placement of the cue ball during a pool tournament); this ability accrues from our early base in simple spatial reasoning.

Spelke and her colleagues (1992) provided evidence for these simple spatial reasoning abilities in infants as young as 2.5 months. These infants looked longer at events where a ball “moved through” a solid barrier to an outside wall (violation of solidity) as compared to a ball that stopped at the inside wall of the barrier. Four-month-olds looked longer at events where a ball “fell through” a solid platform in a vertical fashion (violation of solidity) as compared to a ball that landed on a solid platform.

For gravity violations, the researchers found that 4-month-olds did not show a preference for events where a falling object stopped in mid-air, and 3-month-olds did not show a preference for events where a ball moving horizontally seemed to “jump a gap” in a platform (Spelke et al., 1992). However, Baillargeon (1994) found that 3.5-month-old children looked longer at events where boxes were pushed off of platforms but did not fall (a typically gravity violation). Kim and Spelke (1992) investigated other effects of gravity violations on looking preferences for infants; 7-month-olds looked longer at events where a ball moved up or down a slope with inappropriate acceleration as compared to events with appropriate acceleration, though 5-month-olds did not exhibit a preference.

In investigating infants’ reasoning about object movement based on continuity, Spelke and colleagues (1994) found that infants aged 4-, 6-, and 10-months looked longer at events in which a ball “moved” to a location that was inconsistent to the physical law of continuity. Collectively, the researchers conclude that these findings suggest that infants under 10 months of age (and thus, younger that those in Stage V of the sensorimotor period) have considerable spatial reasoning skills for understanding simple physical events.

Contrary Evidence Based on Preferential-Looking Methods

Questions arise as to the validity of the interpretations made about the spatial reasoning skills of infants based on data from the preferential-looking studies. Cohen, Gilbert, and Brown (1996) provided a task similar to Spelke et al.'s (1992) to infants aged 4-, 8-, and 12-months to investigate their understanding of solidity (an object cannot fall through a solid platform) and gravity (an object cannot stop in mid-air over a hole in a platform). In their experiments, only the 12-month-olds looked longer at the impossible events than the possible events.

Kannass and Oakes (1997) investigated infants' preferences for events consisting of gravitational constraints on moving objects. However, they also investigated the effects of manipulating the habituation trials on infant preference. When no habituation trials were presented, 10- and 16-month-old infants showed no a priori preference for impossible events (e.g. a ball accelerating up an incline) compared to possible events (e.g. a ball decelerating up an incline). Regardless of the type of habituation stimulus provided (possible or impossible), 10-month-olds exhibited a preference for the novel test pattern; 16-month-olds exhibited a preference for a new impossible event when habituated to an impossible event if the new impossible event showed the ball moving in a different direction.

Schilling (2000) further investigated the role of familiarization or habituation trials in looking preference in infants. He presented 4- and 6-month-old infants with screen rotation events where the screen rotated either 180 degrees or 112 degrees. Each age group received seven or 12 familiarization trials of either degree rotation events at familiarization trials. He found that 4-month-olds looked longer at the opposite event during testing (e.g. the 112 degree event when habituated to the 180 degree event) when they received 12 familiarization trials, but that the opposite effect was found when they received seven familiarization trials. Six-month-olds receiving seven familiarization trials did not look longer at either type of test trial. Summarizing these findings from Cohen et al. (1996), Kannass and Oakes (1997), and Schilling (2000), we see that manipulating the number and type of habituation events can affect the outcomes of preferential-looking studies with infants.

Given these contrary findings, we cannot conclude that young infants are capable of spatial reasoning about continuity, gravity, and solidity. A more economical conclusion may be that the infants are capable of recognizing that the impossible event is visually very different from the possible event to which they have been habituated, and/or the novel possible event is visually very similar to the habituated possible event. Though the events seem similar in presentation to the adults except for the possibility of the event, the infants may be preferentially looking at the stimuli based on visual features. Therefore, the infants may not have as accomplished spatial reasoning skills as reported by Spelke et al. (1992) or others (e.g. Baillargeon, 1994; Kim & Spelke, 1992).

Using Active Search and Pointing as Behavioral Responses

More recent research has attempted to replicate the core hypothesis findings, but by utilizing active search and pointing as the required behavioral response given by the young children (Hood, 1995, 1998; Hood, Carey, & Prasada, 2000). In these studies, the event occurs within view, and the child must indicate through pointing or active search the end-point location of an object that has moved into a container or behind an occluder. This requirement does not allow younger infants to participate (they do not yet have the motor coordination to do so); however, if younger infants are capable of basic spatial reasoning, then older infants and preschoolers should be capable of succeeding at these other tasks.

Hood (1992) presented children aged 2 to 4 years with a task that required them to indicate by pointing or searching which container a ball traveled to when falling through an opaque tube. The number of tubes presented during a trial ranged from one to three, and differing diagonal tube positions were given. Only the older children were capable of predicting which container held the fallen ball when diagonal opaque tubes were used; the common incorrect response was to choose the container directly below the dropping point. When transparent tubes provided the paths, all children were able to follow the path of the moving ball and indicate its end-point location, even with diagonal tubes. However, this knowledge did not transfer back to the opaque tube condition. Hood concluded that only older preschool children were capable of predicting object movement and end-point locations based on gravity and the solidity of the tubes, and thus the spatial reasoning abilities proposed by infant preferential-looking method studies should be questioned.

Hood (1995) presented a modified version of this task to another group of 2-year-olds. In this experiment, the child was required to watch a video image of a ball dropping down an opaque tube and then point to the container on the screen that should be the end-point location of the ball. Again, the 2-year-olds pointed to the container that resided directly below the starting point of the dropped ball even though the opaque tube was arranged diagonally from the start point to another container.

Hood, Carey, and Prasada (2000) investigated spatial reasoning for solidity in a modified version of the tasks employed by Spelke et al. (1992). Two-year-old children were presented with trials that required them to search behind curtained windows for objects that fell downwards or rolled horizontally behind a screen. Trials consisted of instances where there was no shelf separating the two windows (i.e. the object could fall to the bottom or roll to the other side behind the screen), or a shelf was placed between the two windows in the child's presence (i.e. the object could only fall to the shelf behind the top window or roll to the barrier beside the first window). Two-year-olds showed no preference for searching the windows based on the presence or absence of the shelf. Hood et al. (2000) concluded that 2-year-old children were not capable of spatial reasoning for solidity issues, and they question the interpretations made by those who subscribe to the core knowledge hypothesis.

Variations in Testing Design

How can we account for the differing results across the reviewed studies? The various results could be due to differing task designs. The tasks used to investigate spatial reasoning in young children differ on visual presentation and required behavioral responses. The use of preferential-looking paradigm requires only looking as a response by the infant, and the presentation of an event allows viewing the progression of the moving object until it comes within a small distance of the support or barrier at which time an occluder appears to allow the experimenter to place the object in the possible or impossible location (e.g. Baillargeion, 1994; Spelke et al., 1992). The tube tasks (Hood, 1995, 1998) require prediction of the end-point in trials where the object cannot be viewed as it falls (when tubes are opaque), and the child is required to point or tell the experimenter the expected end-point. These tasks require the child's working memory and attention abilities to follow an event through time.

Secondly, many preferential-looking experiments use habituation trials of possible events before presentation of probe and control trials to determine preference, whereas active search or pointing tasks use limited familiarization trials. Habituation trials may be predisposing the participant to react by looking longer at the impossible events in Spelke et al.'s tasks (1992) regardless of any early knowledge of physical laws. Elimination of habituation trials in preferential-looking tasks and opaque presentations in tube tasks may help reduce the differences between the two testing paradigms. However, Hood and colleagues (2000) used presentations that afforded extended visual contact with the moving object and eliminated habituation in the procedure, and the results continued to indicate that older infants (2 years of age) were unable to indicate end-point locations based on solidity constraints. Additionally, Hood, Cole-Davies, and Dias (2003) presented 2.5- and 3-year-old children with an invisible displacement task (a ball rolling behind a row of four doors) in which they measured looking time for possible and impossible events as well as first choice in a search format; this provided a means in which to test the looking method and search method with the same apparatus. Though the younger and older children looked longer at impossible events compared to possible events, only the older children were successful in searching for the ball. If we cannot reconcile the differing results as reflecting methodology simply based on behavioral responses and visual presentation alone, then how can we bridge the gap between these studies? An examination of comparative studies may help answer this question.

Nonhuman Primate Evidence

Human infants are not the only primate species that show inconsistent results in spatial reasoning tasks. Cotton-top tamarins (*Saguinus oedipus*) also search incorrectly for objects falling through opaque tubes as presented in Hood's tasks (Hood, Hauser, Anderson, & Santos, 1999) and for objects rolling down occluded ramps (Hauser, Williams, Kralik, & Moskovitz, 2001). However, as noted by Santos and Hauser (2002), the tamarins look longer at impossible events than possible events when the same apparatus is used for preferential-looking tasks. Rhesus macaques (*Macaca mulatta*) exhibit inconsistent searches in similar tasks (Hood, 2001), but look longer at impossible events compared to possible events in the preferential-looking version of the same tasks (Santos & Hauser, 2002). These findings are consistent with those found with infants and preschoolers, which is surprising considering that all nonhuman primates that are tested in these studies are adults, and thus should be able to show preferential looking and correct searching during testing.

However, evidence from spatial reasoning studies with tufted capuchin monkeys (*Cebus apella*) may help us bridge the gap between the preferential-looking method and active search studies for all primate species (including humans). Cummins-Sebree and Fragaszy (2000) presented capuchins with modified versions of the tasks used by Spelke et al. (1992) to investigate their ability to predict end-point locations of moving objects based on continuity, solidity, and gravity constraints. For the continuity task, the capuchins were required to indicate by hand placement the end-point of a moving ball across a smooth platform. Training trials were paths that originated in front of the experimenter and ended perpendicular to the subject; probe trials consisted of diagonal paths. For the solidity tasks, hand placement was used to indicate the potential end-point of a circling object across a platform; training trials contained no barrier, whereas the probe trials contained a solid barrier that blocked one direction in which the ball could travel. In the gravity task, training trials involved a ball moving toward one direction on a platform, and probe trials involved a ball moving toward a hole in the platform. Three of six capuchins correctly chose the end-point location for the continuity task at above chance levels. One of six capuchins correctly chose the end-point location for the solidity task at above chance levels, whereas two others chose correctly more often than choosing incorrectly. However, none of the six capuchins correctly chose the end-point location in the gravity task. In fact, the capuchins reliably chose the end-point that would be correct if there had been no hole in the platform on which the ball was traveling.

When considering these results, Cummins-Sebree and Fragaszy (2000) proposed that the difference in results for the solidity and gravity tasks may be due to the presentation of the test trials. In the solidity task, the ball was moved in a small circle to indicate the beginning of the trial; the subject was then required to indicate where the ball *could* move based on the presence or absence of a solid barrier. In the gravity task, the ball was moved toward the hole and stopped when it reached the edge of the platform; the subject was then required to indicate where the ball *would* go. Because of this, Cummins-Sebree and Fragaszy (2000) presented the capuchins with a modified version of the gravity task where probe trials resembled the solidity probe trials, except for the presence of a hole instead of a barrier. Though none of the six subjects chose correctly significantly above chance, they chose more correctly than incorrectly in this modified version of the task.

Spatial Reasoning: Recognition, Production, and Prediction

Based on the capuchin results and the varying results found in spatial reasoning experiments with young children and other nonhuman primates, I propose the existence of three differing spatial reasoning processes: recognition, production, and prediction. These three processes require attention to different spatial information and abilities to use that information, and they may be evident in different studies on spatial reasoning.

Recognition is the ability to attend to the physical features of events unfolding within view and to differentiate the event from those that have been previously seen. Individuals with this ability can perceive the visual differences between two events, but may not be able to understand the great issues involved (i.e. physical laws). The results of preferential-looking tasks that indicate that young infants respond to impossible events based on core knowledge of physical laws may be due to the infants' ability to recognize the visual differences between the habituation trial and the impossible event and respond accordingly. The findings by Kannass and Oakes (1997) and Schilling (2000) wherein preferences were generated based on the number and type of habituation events used support this, and this issue has become quite an important one in the discussion of infant spatial reasoning skills (Hood, 2001). For the capuchins (Cummins-Sebree & Fragaszy, 2000), their ability to indicate where an object could go based on recognizing the differences between one side versus another (hole or barrier versus a smooth platform) could be further evidence.

Production requires detecting the physical features of the situation and object at hand and the producing, through a motor act, the solution to a problem. The development of this ability involves learning what features of the environment are important to pay attention to so as to produce successful problem solving. One must employ actions on the environment to pull out important information concerning what one can and cannot do based on the characteristics of the environment (Gibson & Pick, 2000). The preferential-looking tasks and the tube tasks do not directly assess this ability. However, this may be the ability that is being developed during the sensorimotor period as described by Piaget (1954). As for evidence from capuchin studies, Cummins (1999) found that capuchins could learn to avoid a barrier and a hole (with the barrier being attended to much earlier than the hole) in a platform when using a hoe tool to retrieve a treat. This ability was not evident at the start of testing, but developed with additional chances to produce the correct motor responses (i.e. moving the treat around the hole or barrier).

Prediction is the ability to detect the physical features of the situation at hand and predict what can happen next in an event sequence in which one does not control. An individual capable of prediction can indicate where an object will travel to without seeing an extended presentation of the moving object. This may be the ability that is being exhibited by older children in the tube tasks and is lacking in the younger children (Hood, 1995; Hood, 1998); the older children can detect the physical layout of the presentation and then infer the end-point of the ball, whereas the younger children are unable to infer the path taken by the ball when they cannot see its movement. For the capuchins, their failure to indicate where the ball would go once it came to a hole in the platform could reflect a failure to predict the eventual downward movement through the hole (Cummins-Sebree & Fragaszy, 2000).

Each of these processes involves varying abilities to detect affordances of the present situation and to respond to moving objects accordingly (Gibson, 1979). Detecting affordances requires perceiving the structures of objects and surfaces and acting on those based on the relationship between the physical features of those objects and surfaces and the actions the individual is capable of exhibiting (Gibson, 1988). Each process requires perceiving particular information and exhibiting the appropriate behavioral response one is capable of performing. For recognition, one must only detect physical differences to respond with "surprise" at an event that looks different from what has been seen in the past. For production, one must detect the physical features of the situation and produce appropriate motor actions to solve a problem. To be capable of prediction, one must perceive the physical features of the situation, infer the next event in an event sequence, and indicate that to another individual, either through vocalization or motoric action (e.g. pointing, searching).

E. J. Gibson and Pick (2000) outlined a theory of perceptual development based on the ecological theory of perception developed by J. J. Gibson (1979). One's ability to detect affordances in the environment is contingent upon the individual's ability to act on that environment. For example, crawlers and experience walkers perceive different surfaces as traversable (Adolph, 1997). When crawlers begin walking, they attempt to do so on surfaces they have previously crawled on successfully. However, they discover quickly that a slope that allows crawling does not necessarily allow walking (i.e. it is too steep). Walkers learn to detect new affordances (what makes a traversable surface) only after implementing their newly developed action pattern.

Based on the findings reviewed here, I propose that there is a developmental trajectory for these three spatial reasoning processes, and the occurrence of the trajectory can be illustrated through the ecological approach to perceptual development. Based on Spelke et al. (1992) and Baillargeon's (1994) work, I propose that we develop our recognition ability early in life. As we become more agile and develop the ability for more motoric actions, we develop our ability for production; we do this through active exploration of the objects in our environment and how they behave based on actions done on them (i.e. Piaget, 1954). Through our increased ability to produce appropriate responses to solve spatial problems, we develop our prediction ability so that we can now infer the endings of event sequences, such as the end-point location of an unseen moving ball based on the contours of its container (i.e. Hood, 1995). The purpose of this study was to investigate the existence of these three processes in young children, ages 24 to 44 months and delineate the development of these processes accordingly.