INTEGRATING A PHYSICAL CONSTRAINT INTO A STATISTICAL INFERENCE MECHANISM BY 11-MONTH-OLD INFANTS

by

Stephanie Denison

B.A., The University of British Columbia, 2006

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF ARTS

in

The Faculty of Graduate Studies

(Psychology)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

August 2009

ABSTRACT

The majority of research on cognitive development focuses on either early-emerging domain-specific knowledge or domain-general learning mechanisms; however little research examines how these sources of knowledge might interact. Previous research suggests that 8-month-old infants can make inferences from samples to populations and 11-month-old infants can integrate psychological knowledge into this mechanism (Xu & Garcia, 2008; Xu & Denison, in press). Here we asked whether infants can integrate a physical constraint, namely, a violation of cohesion into this statistical inference mechanism. Infants succeeded at this in two ways: First, they were able to override statistical information in favor of domain-specific knowledge, reasoning that in some cases a physical constraint is more informative than probabilistic information. Second, they were able to integrate the constraint into the mechanism by using it to exclude a set of objects and then computing probabilities over two remaining sets.

TABLE OF CONTENTS

List of Figures

Acknowledgements

I thank my supervisor Fei Xu for her invaluable guidance regarding the design of the current experiments and the preparation of this manuscript. Thanks also to my committee members Janet Werker and Susan Birch for their many thoughtful insights. I also thank Vanessa Waechtler, Sarah Heller and Maria Ho for their help in running participants and to the members of the Baby Cognition Lab for their constructive comments on the design of these experiments. Finally, thanks to the babies and their parents for their participation.

1. Introduction

The focus of research on cognitive development in recent years highlights both early knowledge and concepts in infancy as well as the importance of learning and inferential mechanisms. However, the majority of research in the field to date has explored either early knowledge or learning mechanisms but has not investigated how these two sources of knowledge interact. The present paper explores whether or not young infants can integrate domain-specific knowledge into a domain-general learning mechanism.

This introductory section includes four components. I begin with a review of the literature suggesting that young infants possess a rich understanding of physical knowledge. The second section details a number of inferential mechanisms and learning mechanisms available to infants and young children. In the third section, I discuss the relatively small number of experiments that investigate the interaction between domain-specific knowledge and learning mechanisms in adults, children, and infants. Finally, I conclude the introduction by outlining a set of three experiments that investigate whether 11-month-old infants can integrate substantive domain-specific knowledge into a statistical inference mechanism. More specifically I address the question of whether infants can integrate a physical constraint into a statistical inference mechanism in order to reason about the relationships between samples and populations.

Physical Knowledge in Infancy

The literature exploring early knowledge reveals that infants and young children possess domain-specific knowledge to facilitate reasoning about agents, number, objects and causality (e.g., Baillargeon, 2004; Carey, in press; Gergely, Nadasdy, Csibra, & Biro, 1995; Leslie & Keeble, 1987; Spelke, 1990; and Xu & Spelke, 2000). This early knowledge could be innately given or acquired (Carey & Spelke, 1994; Cosmides & Tooby, 1994; Elman, Bates, Johnson, Karmiloff-Smith, 1996; Hirschfeld & Gelman, 1994; Karmiloff-Smith, 1992; Gopnik & Meltzoff, 1996; Keil, 1989; Pinker, 1994; Smith, 2001; Spelke, 1994). Regardless of how infants obtain this knowledge, much evidence corroborates its existence—showing that infants possess a great deal of knowledge regarding multiple domains very early in development.

One domain that has been extensively investigated in early infancy is that of physical knowledge, and evidence of this knowledge is apparent in infants as young as 2 months of age (Baillargeon, 2004, 2008; Spelke, 1990, 1994; and Spelke, Breinlinger, Macomber, and Jacobson, 1992). According to Spelke's core knowledge thesis, when reasoning about physical events, infants use three basic principles— cohesion, continuity, and contact— to make inferences about the behaviour of hidden objects (Spelke, 1994). A great deal of research has been conducted to explore at what ages infants begin to reason in accord with these core principles and to what extent this understanding facilitates their interpretation of more complex physical events.

The principle of cohesion, which states that objects should move as connected, bounded units (Spelke, 1994) has been explored in infants using looking-time measures (Spelke & Born, 1983; Spelke & Van de Walle, 1995). In an experiment conducted by Spelke and Born (1983), 3-month-old infants were shown a display with a conical object placed on an empty stage. Infants in an experimental condition were habituated to trials where an experimenter tapped the top of the object and the object remained stationary. Infants in a baseline condition were not familiarized to the object prior to test trials. Then, on alternating test trials, all infants saw a hand grasp the top of the object and either lift the entire object as a whole or lift only the top half of the object. Infants in the experimental group looked longer at the event where only the top half of the object was lifted, whereas infants in the baseline group did not look longer at either outcome. This suggests that infants in the experimental group (who had been familiarized to the object as one coherent unit) inferred that the object was a single whole that should not spontaneously break in half when lifted. This evidence suggests that by approximately 3 months of age, infants are capable of reasoning about physical objects in accord with cohesion.

The second principle proposed by Spelke to be represented in infants' system of core knowledge of physical events is that of continuity—objects move on connected paths through space and time (Spelke, 1994). The principle of continuity encompasses both the understanding that objects move on connected paths as well as an understanding of solidity (i.e., that two objects cannot occupy the same space simultaneously). Spelke, Kestenbaum, Simons, and Wein (1995) tested whether 4-month-old infants expect objects to move in accord with continuity. They showed two groups of infants different events involving two occluders aligned side-by-side on a stage, separated by a gap. They familiarized one group of infants to an event where a rod began on the left side of the stage and moved across the stage—it traveled behind the first occluder, appeared in the gap between the two occluders, traveled behind the second occluder and then appeared on the opposite side of the second occluder. They familiarized a second group of infants to an event where the rod began on the left side of the stage and moved across the stage and traveled behind the first occluder. Then a second rod appeared from the right side of the second occluder without any rods appearing in the gap between the two occluders. All infants were then shown the same displays in plain view, one event involving a single rod moving continuously across the stage and one event involving two rods moving discontinuously across the stage. Infants' looking behaviour suggested that those in the first group had inferred that the scene they were habituated to consisted of one object moving continuously across the stage. Those in the second group had inferred that the scene they were habituated to consisted of two objects moving discontinuously.

Spelke, Breinlinger, Macomber, and Jacobson (1992) tested whether or not 4-month-old infants could detect violations of solidity when reasoning about the actions of a hidden ball. Infants were habituated to a scene where a ball was dropped behind an occluder and the occluder was then lifted to reveal the ball on the floor. On test trials, infants saw the ball being dropped behind the occluder and when the occluder was removed, they saw, on alternating trials, either the ball sitting on a platform that had been placed above the floor or the ball sitting on the floor underneath the platform. Infants dishabituated to the physically impossible event of the ball sitting below the platform. This pattern of looking suggests that infants expected the event that occurred behind the occluder during habituation to consist of the ball falling to an impermeable floor or platform. Therefore, when the ball appeared to have passed through the solid platform, this violated infants' knowledge of solidity. Taken together, the results from Spelke and colleagues (1992, 1995) suggest that infants as young as 4 months of age are able to reason about objects in accord with the principle of continuity—they believe that individual objects move on continuous paths, and that solid objects cannot pass through other solid objects.

Along with a basic understanding of cohesion and continuity, researchers have also found that infants expect objects to behave in accord with the principle of contact—objects affect one another's movement if and only if they touch (Needham, 1999; Spelke, 1994). For example, Ball (1973) conducted an experiment investigating whether infants' knowledge of physical objects includes accurate reasoning regarding the contact principle. To test this understanding, infants of a variety of ages were familiarized to displays where one object sat in plain view next to an occluder and another object was placed such that half of the object was visible and the other half was hidden behind the occluder. Infants watched the first object move behind the occluder in line with the second object's position and then saw the second object move out the other side of the occcluder. Next, infants were shown two alternating events in plain view. Infants dishabituated to events where the first object did not come in contact with the second object to initiate movement, but did not dishabituate to events where the first object did make contact with the second object. These results suggest that, during habituation, infants inferred that the first object made contact with the second object behind the occluder. This finding has been replicated a number of times, indicating that beginning at about 3 month of age, infants expect objects to affect one another's motion only when they make contact, and that inanimate objects cannot act on one another at a distance (Leslie, 1988, 1994; Spelke & Van de Walle, 1993).

In addition to the evidence suggesting that infants have expectations about objects in agreement with these three core principles, researchers have also demonstrated that they can use this fundamental knowledge to reason about more complex physical events. These include occlusion, containment, and covering events. Infants' understanding of these events has been documented at as young as 2.5-months of age and continues to develop throughout the first year of life. For example, Aguiar and Baillargeon (1999) demonstrated that 2.5-month-old infants can represent simple occlusion events. They showed infants an event where a mouse disappeared behind a screen and then reappeared from behind another disconnected screen. Infants' looking patterns suggested that this type of event violated their expectations. In a containment event, Hespos and Baillargeon (2001) showed 2.5-month-old infants' an object being lowered inside a container and then the container being moved forward and to the side. Infants' expectations were violated when the object remained in the place where the container originally sat. Finally, in a covering event, 2.5-month-old infants saw a toy duck sitting on one side of a platform and then saw a cover being placed over the duck. When the cover was slid to the right and lifted to reveal no duck, infants' expectations appeared to be violated (Wang, Baillargeon, & Paterson, 2005).

The above evidence indicates that infants possess early knowledge that facilitates reasoning about events concerning the domain of physical objects. In sum, infants appear to have an early understanding of a variety of physical properties relevant to objects. At the most fundamental level, young infants understand that the actions of objects should adhere to the principles of contact, continuity, and cohesion. Infants can use this information to reason about physical events involving occlusion, containment, and covering.

Learning and Inference Mechanisms

Parallel to the research conducted on early domain-specific knowledge, researchers have also focused on infants' use of powerful learning mechanisms that allow them to compute various statistics over input received from the environment (e.g., Aslin, Saffran, & Newport, 1998; Gerken, 2006; Gomez, 2002; Gopnik, Glymour, Sobel, Schulz, Kushnir, & Danks, 2004; Marcus, Vijayan, Rao & Vishton, 1999; Maye, Werker, & Gerken, 2002; Saffran, Newport, & Aslin, 1996; Smith, Jones, Landau, Gershkoff-Stowe, & Samuelson, 2002; Sobel & Kirkham, 2006, 2007; Teglas, Girotto, Gonzalez, & Bonatti, 2007; Xu & Garcia, 2008).

A number of statistical learning and inference mechanisms in infants have been studied in both linguistic and non-linguistic contexts. In the domain of language, researchers have investigated what kinds of statistical mechanisms infants are capable of using when acquiring their native languages (Maye et al, 2002; Aslin et. al, 1998; Saffran et. al, 1996; Marcus et. al, 1999).

When acquiring language, infants must begin by discriminating the perceptual properties of their language. A possible strategy for discriminating the relevant acoustic dimensions in a language is to attend to the distributional properties of the sounds in that language. Maye et al. (2002) proposed that one mechanism infants can use to discriminate phonetic categories is to attend to the shapes of the distributions of different acoustic dimensions in phonetic categories. For example, differences in the acoustic dimension of voice-onset-time (VOT) are important for distinguishing the phonemes /b/ and /p/. This results in tokens of /b/ and /p/ forming a bimodal distribution such that most exemplars cluster at opposite ends of the distribution, centered on /b/ and /p/. In contrast, if an acoustic property is relatively unimportant in distinguishing phoneme categories, the distribution will be unimodal, clustering in the center of the distribution for that property. They tested whether 6- and 8-month-old infants could identify VOT as relevant for distinguishing phonemes using a preferential looking-time measure. They familiarized two separate groups of infants to either bimodal or unimodal distributions and then played test trials for the infants with either alternating (e.g., /p/ and /b/) or non-alternating sounds. They found that infants in the bimodal condition looked longer during non-alternating trials, whereas infants in the unimodal condition showed no preference for either trial type. This finding suggests that infants can use distributional information as a cue to the phonetic structure of their native language by highlighting which acoustic properties are relevant for discriminating phonemes.

Another problem that young infants must solve early in development in order to acquire language is that of word segmentation—they must be able to parse speech into sets of syllables separated by word boundaries. A number of researchers suggest that this could be achieved by a mechanism that allows infants to compute transitional probabilities across syllables (Aslin et al, 1998; Saffran et al, 1996; Saffran, Newport, Aslin, & Tunick, 1997).

Aslin et al. (1998) tested whether 8-month-old infants were capable of using transitional probabilities to segment speech. Transitional probabilities are a type of conditional probability calculated such that the probability of Y given X is equal to the frequency of X and Y divided by the frequency of X. For example, when one considers the word "baby", *bay* is a syllable that will be followed by *bi* 100% of the time in the artificial language because it occurs within a word. Of course within sentences, the word "baby" can be followed by a vast number of words (e.g., cries, crawls, eats, etc.). Therefore, syllable pairs that appear within words occur more frequently than syllable pairs that occur across word boundaries.

In the artificial language used by Aslin et al. (1998), infants were exposed to a continuous speech stream consisting of 3-syallable non-sense words presented in random order. The stream was arranged such that syllables heard within words had transitional probabilities equal to 1.0 and syllable pairs between words had transitional probabilities that were much lower (they controlled for frequency by matching words and part-words in frequency but not transitional probabilities). The experimenter then assessed infants' learning by presenting them with syllable strings of either "words" from the original speech stream or “part-words” consisting of syllables that were present across word boundaries during familiarization but were not contained within individual words. Infants showed a novelty preference for the “part-words”, indicating that they had tracked the transitional probabilities of the syllables in the speech stream and used this information to segment words.

These results, coupled with the results from Maye et al. (2002), suggest that infants are capable of using statistical information including the shapes of distributions of acoustic properties as well as sequential statistics as a cue to word boundaries to acquire language. In conjunction with these findings, a number of other experiments have revealed additional inferential learning mechanisms that human infants can use in language acquisition. These experiments demonstrate that infants and adults can use rule-learning mechanisms to detect and generalize abstract statistical rules from linguistic input (See Gomez, 2002; Gerken, 2006; Marcus, Vijayan, Rao & Vishton, 1999).

Although much of the work on learning mechanisms in humans focuses on language acquisition, there also exists a body of research on learning mechanisms in infants, children, and adults outside of the linguistic context. In other domains, researchers have examined the statistical mechanisms involved in prototype abstraction, causal reasoning, and probabilistic inference (Gopnik et. al, 2004; Quinn, 1987; Sobel & Kirkham, 2006, 2007; Teglas et. al, 2007; Xu & Garcia, 2008; Younger, 1993).

Some earlier studies suggest that young infants are able to extract abstract category information (i.e., prototypes) from visual displays (Quinn, 1987; Young, 1993). More recently, researchers have begun examining causal inference mechanisms, mostly in young children but also in infants (Gopnik et al, 2004; Gopnik, Sobel, Schulz, & Glymour, 2001;

Sobel & Kirkham, 2006, 2007). These mechanisms are of particular interest, as they appear to allow young learners to make rich inferences from relatively sparse amounts of data.

For example, Gopnik, Sobel, Schulz, and Glymour (2001) investigated the nature of a possible causal inference mechanism in 2-, 3-, and 4-year-old children. An experimenter taught children a novel causal property using a novel toy called a Blicket detector, which activates when Blickets are placed on its surface. Children were then shown one of two events, one-cause events or two-cause events. In one-cause events, the experimenter brought the detector and two blocks onto the table. He then demonstrated that the first block activated the detector but the second block did not. Next, the experimenter placed both blocks simultaneously on the detector and the machine activated. Finally, he asked the children which blocks were Blickets. In two-cause events, children saw that the first block activated the detector three times and the second block did not activate the detector the first time it was placed on the machine but then activated the machine the next two times it was placed on the machine's surface. The experimenter then asked the children which blocks were Blickets. They found that, in the one-cause events, children most often stated that the first but not the second block was a Blicket. This suggests that, for children, the causal powers of the first block "screened off" any possible causal power that might be inferred to exist within the second block due to the fact that the second block only exhibited causal power when the first block was present. Conversely, in the two-cause events, children most often stated that both blocks were Blickets. This suggests that 2- to 4-year-old children can use causal reasoning to make inferences about novel causal properties.

Two other learning mechanisms that appear to be available to young infants outside of the domain of language may guide their probabilistic inferences. First, in an experiment investigating single-event probabilistic reasoning in 12-month-olds, Teglas et al. (2007) found that infants could reason about probability in a violation of expectation looking-time paradigm. Infants saw a computer screen displaying a set of four objects—three blue objects that were identical in shape and one yellow object that was different in shape—moving around in a lottery machine. On alternating trials, infants saw either one of the identically shaped blue objects or the differently shaped yellow object fall from the machine. Infants looked longer when the less probable object (the different object) exited the machine. This result indicates that 12-month-old infants can use a probabilistic inference mechanism to make inferences about the likelihood of a future single event.

Second, Xu and Garcia (2008) conducted a series of experiments investigating a statistical inference mechanism involving large populations and multi-object samples. They tested 8-month-old infants in a looking time paradigm. In this set of experiments, infants were familiarized to two possible populations—a large population of mostly red Ping-pong balls and just a few white balls and a large population of mostly white Ping-pong balls and just a few red balls. On alternating test trials, infants saw a sample of either four red balls and one white ball or four white balls and one red ball being drawn from a covered box. The population of, for example mostly red Ping-pong balls, was then revealed. Infants looked longer at the less probable sample of four white and one red ball, indicating that this outcome violated their understanding of probability. This finding suggests that infants can use statistical information to make inferences about the composition of a population given data obtained from a small sample.

Integrating Domain-Specific Knowledge and Domain-General Mechanisms

Given that young learners possess both substantive domain-specific knowledge and statistical inference and learning mechanisms, the question arises as to whether infants can integrate these knowledge sources. It is possible that infants are not capable of integrating domain-specific knowledge with domain-general learning mechanisms and that these systems develop separately early in life. However, it is also possible that infants' initial conceptual knowledge interacts with learning mechanisms in informative ways, allowing them to use both types of knowledge simultaneously. Due to the fact that most infancy research investigates either initial knowledge or learning mechanisms, we do not have a great deal of evidence to suggest which of these alternatives represents the true state of affairs. However, some evidence exists with older children and adults suggesting that they can make inferences using both domain-specific and domain-general knowledge (Kushnir & Gopnik, 2007; Newport & Aslin, 2004; Schulz, Bonawitz, and Griffiths, 2007; Schulz & Gopnik, 2004).

In regard to adults, Newport and Aslin (2004) have investigated how domain-specific linguistic knowledge interacts with the mechanism infants and adults can use to compute transitional probabilities. This set of experiments probed whether adults could compute non-adjacent regularities in speech, a necessary task for any language learner— although the problem becomes infinitely challenging when one considers all of the regularities the learner could possibly keep track of (e.g., probability between elements that are one apart, two apart, three apart, etc). They found that the units over which adults can compute non-adjacent dependencies are constrained such that they are highly related to the patterns found in most natural languages. Specifically, they found that participants did not learn non-adjacent dependencies when computing over syllables but did learn the dependencies when computing over phonemic segments (i.e., patterned consonants with intervening vowels or patterned vowels with intervening consonants). This suggests that the initial knowledge humans' posses in the domain of language might interact with the statistical inference mechanism used for segmenting speech.

In addition to this evidence with adults, there exists evidence in the causal reasoning literature that children can integrate information regarding domain-specific initial knowledge and causal learning mechanisms (Schulz, Bonawitz, & Griffiths, 2007; Schulz & Gopnik, 2004). For example, Schulz and colleagues (2007) tested whether 3- and 4-year-olds could integrate constraints from their own domain-specific theories with a domain-general causal inference mechanism. They examined the interplay between theory and evidence in preschool-age children, contrasting preschoolers' predisposition to deny psychological causes for physical events with statistical evidence weighted towards these causes. In their experiment, they presented a relatively large amount of evidence demonstrating that a child being scared was correlated with the child getting a stomachache. This is a cause that children would not normally endorse given their domain-specific theories regarding psychosomatic causes. However, they found that 4-year-olds were in fact swayed by the evidence—if the child consistently got a stomachache after being scared they were willing to accept fear as a valid cause of stomachaches. This suggests that when children are faced with causal evidence that contradicts a domain-specific theory, 4-year-olds will override their initial theory with the statistical information in the input. It is possible that children in this experiment were able to integrate the statistical information and their previous domain-specific knowledge to update their theories.

Along with evidence from adults and preschoolers, two studies suggest that infants can integrate domain-specific knowledge into statistical inference mechanisms (Teglas et. al, 2007; Xu & Denison, in press). In the experiments conducted by Teglas and colleagues (2007, described in the previous section), a control experiment was included which required 12-month-old infants to integrate their knowledge of solidity into a probabilistic inference mechanism. They placed a barrier in the middle of the lottery machine with the three identical objects placed above the barrier and the one different object placed below the barrier. Infants looked longer when one of the identical objects exited the bottom of the machine than when the different object exited, suggesting that they recognized the former as an impossible event. This suggests that 12-month-old infants were able to override probabilistic information and instead use their knowledge of solidity to reason about the likelihood of the test events.

A second experiment investigated whether 11-month-old infants can integrate knowledge regarding agents into a statistical inference mechanism (Xu & Denison, in press). The authors investigated whether infants could reason about psychological constraints, namely an agent's goal and visual access, in the context of Xu and Garcia's (2008) statistical inference task. The procedure was similar to Xu and Garcia (2008), except that the experimenter conveyed a goal to the infants of obtaining one colour of balls over the other. For example, during a familiarization phase, the experimenter demonstrated to the infants that she had a goal of only removing white balls from a container. In one condition, during the test phase, the experimenter looked directly into the box while removing a sample of either five white or five red balls from the box. This visual access indicated that the experimenter was not sampling randomly from the box. In another condition, the experimenter removed the same samples but was blindfolded during test trials. This lack of visual access indicated that, although the experimenter had a goal of obtaining the white balls, she had no means of acting on this goal and thus was forced to draw a random sample from the box. Infants in both conditions were then shown a population of mostly red balls and just a few white balls. Infants in the “ visual access” condition looked longer at the sample of five red balls, presumably because this was inconsistent with the experimenter’s goal. Infants in the “blindfolding” condition looked longer at the sample of five white balls, presumably because this violated the probability of a random sample. These findings suggest that infants were able to integrate domain-specific knowledge regarding agents into a statistical inference mechanism in a meaningful way. They overrode probabilistic information in favor of domain-specific knowledge and vice versa, under the appropriate sampling conditions.

Although these experiments provide preliminary evidence that infants can integrate domain-specific knowledge and statistical inference mechanisms, the findings leave open a number of questions. First, are infants able to integrate other types of domain-specific knowledge? Given the vast amount of literature showcasing infants’ physical knowledge, and the findings of Teglas et al. (2007) regarding solidity violations, it seems likely that this might be a system of knowledge they would succeed at integrating into a statistical inference mechanism. Second, both Teglas et al. (2007) and Xu and Denison (in press) required infants to integrate a constraint into the statistical inference mechanisms only to the extent that infants had to determine which of the two sources of knowledge they should appeal to when faced with conflicting knowledge regarding substantive domain knowledge and probabilistic information. Can infants integrate initial knowledge and a statistical inference mechanism in a situation that does not allow them to completely disregard one type of knowledge but instead requires them to use both simultaneously while making inferences?

Current Experiments

The current experiments investigate whether 11-month-old infants can integrate a physical constraint, namely a violation of cohesion, into a statistical inference mechanism. This set of experiments will shed light on two questions: In Experiment 1 we ask whether infants can detect a physical constraint using the same methodology as Xu & Garcia (2008) and Xu & Denison (in press). We test the hypothesis that, if infants are sensitive to a violation of cohesion, they will use this information to override the probabilistic information available to them when making inferences from samples to populations. Experiment 2 serves as a replication of earlier findings, as well as a control for an alternative interpretation of the looking-time pattern found in Experiment 1. In Experiment 3, we ask whether infants can integrate knowledge regarding cohesion into this statistical inference mechanism. In this experiment, we show infants a box with three sets of Ping-pong balls and ask whether infants can exclude one of these sets of balls and then compute probabilities over the remaining sets.

2. Experiment 1

Method

Participants

Participants were sixteen 11-month-old infants (8 girls and 8 boys; mean age = 10;26 [months;days], range = 10;16 - 11;14). An additional three infants were tested but not included in the final sample due to experimenter error (1) or not completing the study due to fussiness (2). All participants were recruited from the Greater Vancouver area and parents filled out a consent form at the beginning the experiment. Infants received a T-shirt or bib and a diploma for their participation.

Materials

Ping-Pong Balls. A total of 236 (118 green and 118 red) Ping-pong balls were used. Half of the green balls and half of the red balls had six white Velcro strips glued to them so that there was one strip on each side of each ball (strips were approximately .8cm x 1.5cm). The Velcro signified that one of the sets of balls had the property of being physically constrained or immovable. For the remainder of the Method section, we will describe the condition in which the red balls had Velcro and were therefore immovable or stuck inside the box.

Boxes and Containers. A small white box (17.5cm x 17.5cm x 8cm) without a top, constructed out of foam core was used to hold three green and three red Ping-pong balls; the red balls were glued to the box. Infants were allowed to play with the balls in the box during the Free Play Phase at the beginning of the experiment.

A small Plexiglas container (20cm x 4.5cm x 4cm) was placed at the front left-hand corner of the stage to display the five-ball samples pulled out of the box during test trials.

A slightly longer Plexiglas container (28.5cm x 4.5cm x 4cm) was used during Demonstration Phase 1 (see Procedure section below). This container held six balls, three red (with Velcro) and three green (without Velcro). The red balls were glued to the inside bottom surface of the container. Both of the Plexiglas containers were narrow enough such that the Ping-pong balls lined up in a single row when placed inside.

A white box made from foam core, fabric and Plexiglas (27cm x 16cm x 13cm) was used in Demonstration Phase 2 (see below). Purple fabric (secured to the top of the box by Velcro) covered the front of the box and the fabric could be lifted to reveal the front Plexiglas window. The top of the box had a 13cm x 8cm cutout for the experimenter to reach into to access the balls in the box. This box contained 24 Ping-pong balls, 12 green and 12 red. Each red ball was glued to one of the six inside surfaces of the box.

A 39cm x 34cm x 22cm box constructed out of foam core, fabric, and Plexiglas was used to display the large population outcomes. The box was a white rectangular cube lined with black duct tape around all edges. The inside of the box was divided into three parts with two Plexiglas containers inserted into the front and back of the box, each containing 72 Ping-pong balls, and a center compartment used to hold the samples to be removed from the box during test trials. When viewed from the front or the back, the box appeared to be one large box filled with Ping-pong balls. The front and back of the box were covered with black fabric curtains (secured to the top of the box with Velcro) that could be lifted to reveal the contents of the box through the Plexiglas window. The “mostly red” side of the box contained 60 red and 12 green balls (red : green = 5:1), the “mostly green” side contained the opposite ratio (red : green = 1:5); the red balls were covered with Velcro strips. The top of the box had a

10cm x 24cm cutout covered with two pieces of overlapping spandex that allowed the experimenter to reach into the center compartment of the box.

Apparatus

Testing took place in a quiet room. The room was divided in half by two curtains spanning the width and height of the room. All events were presented on a puppet stage. There was a black curtain on the back of the puppet stage, attached at the top of the stage by Velcro. The experimenter sat behind the stage. When the back curtain was lifted, her upper body and head were visible to the infant and when it was pulled down, the experimenter was no longer visible. The viewable area measured 94cm x 55cm (width x height). The observer watched the infant on a TV monitor in one corner of the testing room and recorded each infant's looking times on a Macintosh iBook using MacXHAB 1.4 (J. Pinto, 2005). The observer was blind to the order of the trials. An Optimus fan set at low speed was located in the back part of the room to muffle sounds from the hallway. The stage was lit; the rest of the room was darkened during the study.

The infant sat approximately 70cm from the stage in a high chair and the parent sat next to the infant facing away from the stage. The parent was instructed not to influence the infant's looking behaviour in any way and to face away from the stage throughout the experiment. Two cameras recorded the session, one focused on the stage to record the procedure, the other focused on the infant's face to record her looking behavior.

Procedure

All infants were tested in a violation-of-expectation looking time paradigm.

Free Play Phase. Infants were first shown the small foam core box with three red (with Velcro) and three green (no Velcro) Ping-pong balls and were permitted to play with them for about 1 minute. The experimenter shook the balls around in the box and she encouraged the infants to try to pick up some balls of each colour, giving them the opportunity to notice that the green balls were easily removed from the box and the red balls were stuck.

Demonstration Phase 1. The experimenter went behind the stage and placed the small transparent container on the front of the stage where it remained throughout the experiment. The experimenter then brought out the longer transparent container, which held three red balls (with Velcro) and three green balls in random order. Next the experimenter drew attention to each of the six balls, one colour first, then the other colour, one ball at a time. She did this by picking up each of the green balls while saying, “Look at this one! See this one?” She then grasped each red ball (which did not move) as if she was trying to pick it up and said, “Look at this one; it’s stuck!” The demonstration lasted for approximately 30 seconds after which the experimenter cleared the container from the stage. The purpose of this phase was to demonstrate that there were two colours of balls in the experiment and that the balls with Velcro were immovable.

Familiarization trials. Each infant received four familiarization trials. On each trial, the experimenter placed the large box on the stage with the front curtain closed. She shook the box back and forth a few times, saying, “What’s in the box?” She then lifted the front cover of the box and lowered the backdrop of the stage while saying “Look, [baby’s name], look!” The observer began timing upon hearing the second, “look”, as this was precisely the moment that the experimenter lifted the cover on the box to reveal the population of either mostly red (with Velcro) or mostly green (without Velcro) balls. Once the backdrop was lowered, the experimenter was no longer visible to the infant. The trial ended when the infant looked away for 2 consecutive seconds. The four trials alternated between the mostly green population and the mostly red population. The large box was removed after each trial and the black curtain was lowered between trials. The Familiarization phase lasted approximately 4 minutes.

These trials were included to familiarize infants to the large box and the objects, as well as to the general procedure of the study. Also, once infants have been exposed to two populations, one of mostly green balls, the other of mostly red, they can use this information during test trials to generate a hypothesis as to which box the experimenter might be sampling from before the population is revealed.

Demonstration Phase 2. The experimenter brought the small white box onto the stage. She lifted the front curtain so the infants could see the balls in the box, flipped the box upside down and then right side up, turned the box side to side and shook it and placed it back on the stage. She reached into the box, her hand visible to the infant through the transparent front window. Next, she picked up one green ball and lifted it to the top of the box but did not bring it outside of the box while saying, “See this one? Look at this!” Then she grasped one of the red balls (which was glued to the inside surface of the box) and said, “Look at this one. It’s stuck!” The experimenter lowered the front curtain on the box. She then lifted the curtain again and said, “Should we do that one more time?” and repeated the entire sequence. Following the second sequence, she removed the box from the stage. This phase lasted approximately 45 seconds.

This phase was included so that infants could see that balls with Velcro were stuck inside the box and did not move even when the box was shaken vigorously. This phase also allowed infants to experience a sampling process that is totally transparent: they see the experimenter’s hand reach into the box to pick up balls. This should make the occluded sampling during test trials less mysterious.

Test trials. Each infant received six test trials. On each trial, the experimenter placed the large box on the stage, with its front curtain closed. She shook the box a few times, closed her eyes, turned her head away, and reached into the box. She pulled out three Ping-pong balls of one color (red or green) and placed them into the small Plexiglas container to her right one at a time. She then repeated this action, pulling out two more balls of the same color. The small Plexiglas container had a total of five balls. The experimenter then lifted the front curtain of the box and lowered the backdrop while saying, “Look, [baby’s name], look!” Just as in the Familiarization trials, the observer began timing upon hearing the second “look”, indicating the exact moment that the population was visible to the infant. The trial ended when the infant looked away for 2 consecutive seconds (see Figure 2.1 for a schematic representation of the test trial procedure). Each infant saw one of the two populations (mostly green or mostly red) on all six test trials. At the end of each trial, the stage was cleared. Test trials lasted approximately 7 minutes.

Design

The population of the box (mostly green or mostly red) used for test trials and the colour of balls with Velcro were fully crossed between infants. Half of the infants saw the mostly red population on all test trials, and the other half saw the mostly green population. Half of the infants who saw the mostly green box on test trials saw green balls with Velcro strips and the other half saw red balls with Velcro strips. The same was true for infants who saw the mostly red box. The order of the familiarization trials (mostly red first or mostly green first), the order during the demonstration phases (red or green first), and the order of the samples on the test trials (red or green first) were counterbalanced across infants.

Predictions

We predicted that the 8 infants who saw red balls with Velcro should look longer at the five red ball sample than the five green ball sample. That is, seeing five balls that were demonstrated to be stuck removed from the box should violate their expectations, as this is an impossible event. For the 8 infants who saw green balls with Velcro the opposite prediction holds. Therefore, regardless of the population of the box as mostly red or mostly green, if infants are sensitive to the physical constraint, they should look longer when the sample consists of five balls with Velcro than when the sample consists of five balls without Velcro.

If, however, infants are unable to recognize and integrate the physical constraint, the proportion of balls in the box should predict their looking behaviour. That is, infants should look longer at a sample that is less probable than one that is more probable given the population of Ping-pong balls in the box. For example, for the eight infants who saw the mostly red box on test trials, a sample of five green balls is less probable than a sample of five red balls, thus infants should look longer at the five red ball sample. [1]

Results

Preliminary analyses found no effects of gender, content of the box (mostly red or mostly green), colour of balls with Velcro, or order of the familiarization trials. Subsequent

[1] The probability of obtaining a sample of 5 green balls from the mostly green box = 60/72 * 59/71 * 58/70 * 57/69 * 56/68 = 0.39034 when sampling without replacement. The probability of obtaining a sample of 5 red balls from the mostly green box = 12/72 * 11/71 * 10/70 * 9/69 * 8/68 = 0.00006 when sampling without replacement. The more probable outcome is 6896 times more likely than the less probable outcome.

analyses collapsed over these variables. A second observer who was unaware of the order of the trials timed the familiarization and test trials. Inter-observer reliability averaged 94%.

An alpha level of .05 was used for all analyses. An analysis of variance (ANOVA) examined the effects of test trial order (expected outcome first or unexpected outcome first), trial pair (1, 2, and 3), and outcome (expected vs. unexpected, *according to whether or not the balls were stuck*). There was a main effect of outcome, $F(1,15) = 15.12$, $p < .05$, effect size (ηp^2) = .33. Infants looked reliably longer at the unexpected outcome (M = 9.08s, SD = 3.91) than the expected outcome (M = 6.47s, SD = 2.40) (see Figure 2.2). There were no other main effects or interactions. Thirteen of sixteen infants looked longer at the unexpected outcome, Wilcoxon sign-ranks test: $z = -3.10$, $p < .05$.

However, there exists an alternative interpretation of the data. Due to the counterbalancing of the experiment, half of the infants saw a mostly red box on test trials and half of the infants saw a mostly green box on test trials. Within this, half of the infants saw green balls with Velcro (i.e., green balls were stuck) and half of the infants saw red balls with Velcro. This counterbalancing results in the creation of two groups: One group of infants, termed the Consistent group (n = 8), saw a probabilistic cue and a constraint cue that should predict the same looking time pattern (e.g., they saw a mostly green box and the red balls had Velcro) thus a sample of five red balls is unexpected both in terms of probability and constraints. The other group of infants, termed the Inconsistent group (n = 8) saw a probabilistic cue and a constraint cue that should predict opposite looking time patterns (e.g., they saw a mostly red box and the red balls had Velcro) thus a sample of five red balls is expected in terms of probability but it is unexpected in terms of the physical constraint—the cues were in conflict. The initial analysis does not control for the possibility that infants were insensitive to the physical constraint and that the infants in the Consistent group drove the entire main effect of outcome. To control for this possibility, we re-analyzed the data with the trials coded according to the probability of obtaining the samples from the box if all balls were free to move (i.e., assuming that infants were reasoning about the proportions of the samples and the box and not the physical constraint). That is, if infants saw a mostly red box, the five red ball sample is more probable than the five green ball sample regardless of which balls had Velcro, if the physical constraint is ignored.

Therefore, we conducted an ANOVA examining the within-subjects effects of trial pair (1, 2, and 3), and outcome (expected vs. unexpected, *according to the proportions of the box*) and the between subjects factor of consistency (Consistent group versus Inconsistent group). We found no main effects of outcome (expected vs. unexpected, *according to the proportions in the box*), trial pair, or consistency. This suggests that infants were not reasoning solely in accord with the probabilistic constraint. More importantly, we found an interaction between outcome and consistency, $F(1,14) = 16.31$, $p < .05$, effect size (ηp^2) = .54 (see Figure 2.3). Separating the Consistent and the Inconsistent groups, we found that in the Consistent group, infants looked reliably longer at unexpected ($M = 9.54$ sec, $SD = 1.30$) than expected ($M = 5.98$ sec, $SD = 1.03$) trials, $F(1,7) = 10.58$, $p < .05$, effect size (ηp^2) = .60, *according to the content of the box*. Seven of eight infants looked longer at the unexpected outcome, Wilcoxon signed-ranks test: $z = -2.38$, $p < .05$. In this group it is impossible to discern whether infants were using the probabilistic cue or the constraint cue. In the Inconsistent group, infants looked reliably longer at expected ($M = 8.63$ sec, $SD = 1.30$) versus unexpected ($M = 6.97$, $SD = 1.03$) trials, $F(1,7) = 5.83$, $p < .05$, effect size (ηp^2) = .45, *according to the content of the box*. Six of eight infants looked longer at the expected outcome, Wilcoxon signed-ranks test: $z = -1.96$, $p = .05$. Therefore, infants in the Inconsistent group looked longer at a sample of, for example, five red balls with Velcro from a mostly red box, suggesting that their expectations were violated when balls that should have been stuck inside the box were removed, regardless of the fact that the probability of obtaining that sample from the population (if the constraint is ignored) was very low. This suggests that the physical constraint overrides the probability calculation in this situation.

Discussion

When infants were given evidence of a physical constraint violation—in this case a violation of cohesion—they looked longer at a sample of five balls that violated cohesion than at a sample of five balls that did not violate this constraint. These results suggest that infants were sensitive to the physical constraint of cohesion in this task. The results from infants in the Inconsistent group (i.e., infants who saw probabilistic information that conflicted with the physical constraint) are particularly striking. These infants were given two conflicting sources of information: First, they were shown that balls with Velcro had the property of being stuck to the inside surfaces of the box. Second, they were shown that a sample of five balls was drawn from a population of Ping-pong balls that contained a ratio of 5 balls with Velcro to 1 ball without. Therefore, the probabilistic information in the population should suggest that a sample of balls with Velcro is more probable than a sample of balls without Velcro. However, the physical constraint information should suggest that the sample of balls without Velcro is much more likely. It appears that infants properly integrated these two sources of information, overriding probabilistic information in favor of substantive domain-specific knowledge.

An alternative interpretation of infants' looking time patterns is possible, however. Because the unexpected outcome always involved five balls covered with Velcro, infants may have simply been looking longer at balls with Velcro than balls without Velcro. Therefore, in Experiment 2, we attempt to control for this possibility. Experiment 2 also provides a replication of Xu & Garcia (2008) and Xu & Denison (2009). Those experiments tested whether 8- and 11-month-old infants, respectively, could make inferences about populations given the composition of a sample. Due to the fact that this is a relatively new line of research, it is important to replicate earlier findings.

Figure 2.1. Schematic Representation of the Procedure of Experiment 1.

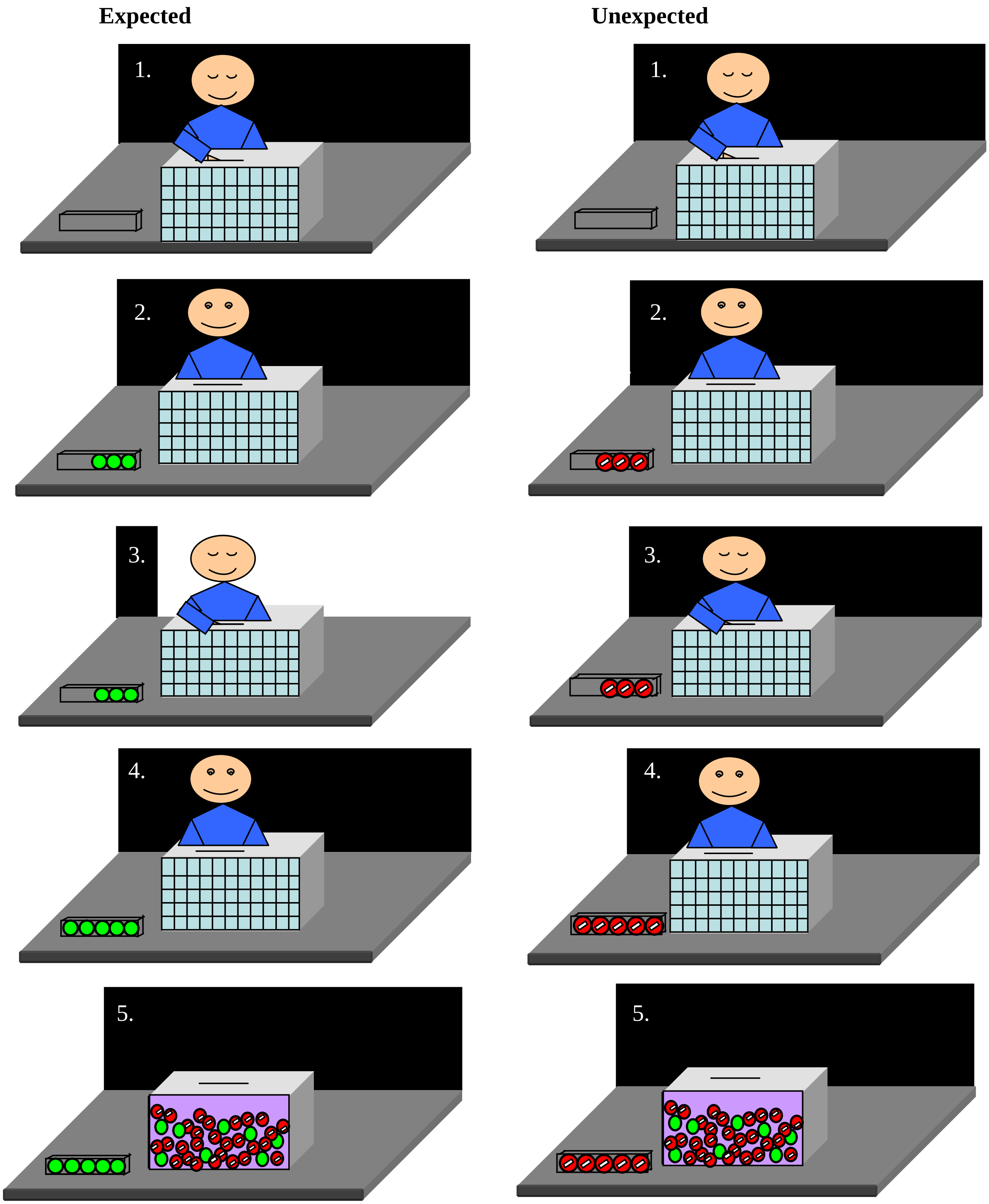


Figure 2.2. Average Looking Time at Expected and Unexpected Outcomes in Experiment 1.

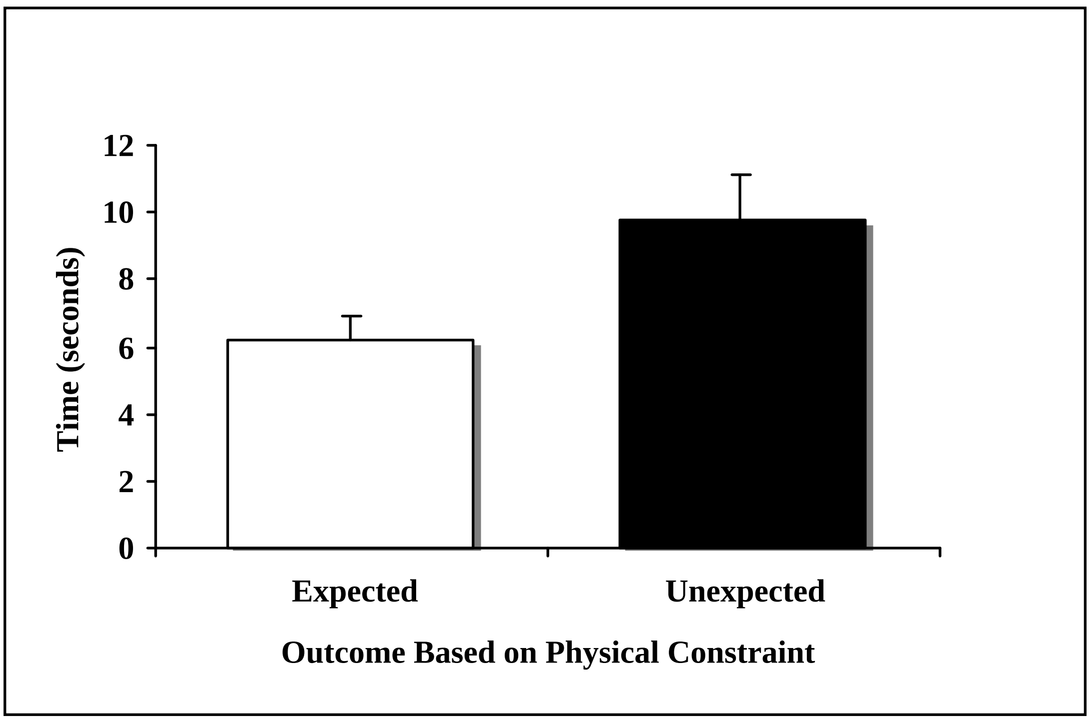
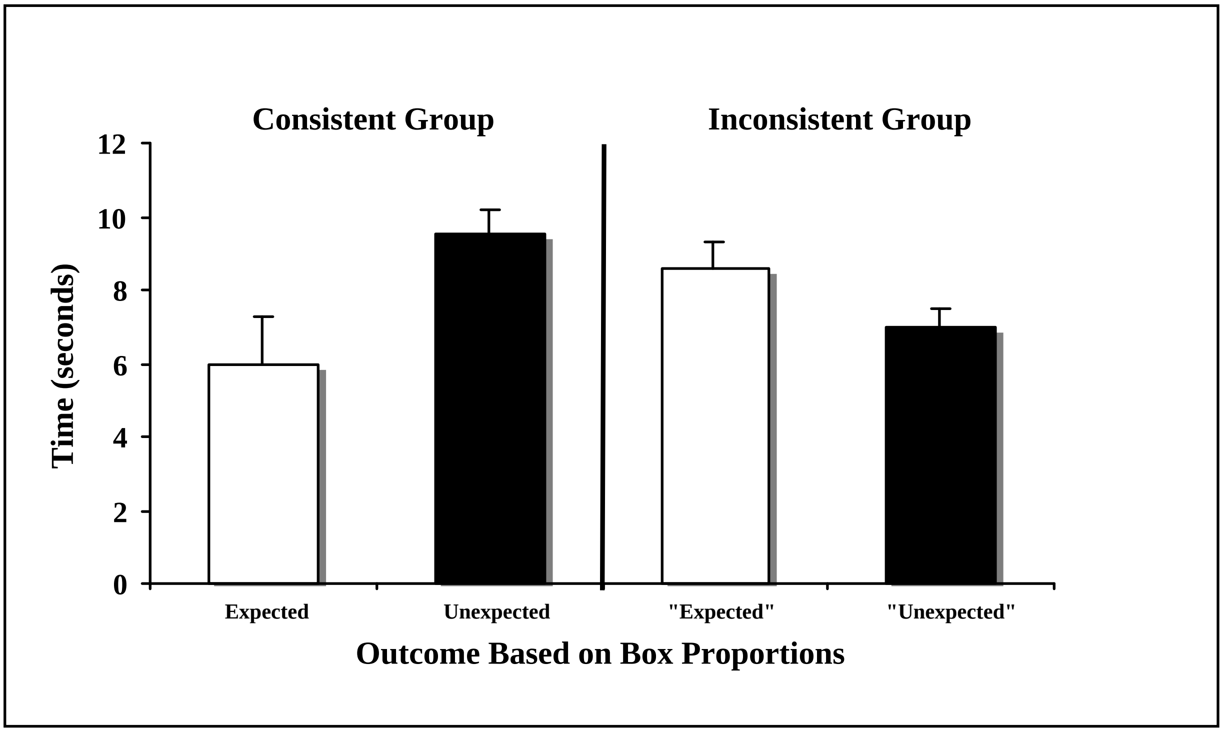


Figure 2.3. Average Looking Time at Expected and Unexpected Outcomes in Experiment 1 Coded According to Box Proportions for the Consistent Group and the Inconsistent Group.



3. Experiment 2

Method

Participants

Participants were sixteen 11-month-old infants (8 girls and 8 boys; mean age = 11;1 [months;days], range = 10;16 to 11;14). An additional four infants were tested but not included in the final sample due to experimenter error (2) or not completing the study due to fussiness (2). All participants were recruited from the Greater Vancouver area and parents filled out a consent form at the beginning the experiment. Infants received a T-shirt or bib and a diploma for their participation.

Materials and Apparatus

All materials were the same as those used in Experiment 1 with one critical exception: Balls with Velcro were not glued to any boxes or containers. The apparatus was the same as Experiment 1.

Procedure

The procedure was the same as in Experiment 1 except that the balls with Velcro were not demonstrated to be stuck. Infants saw that all balls moved freely.

Free Play Period. At the beginning of the experiment, the experimenter showed infants the small white box with three red (with Velcro) and three green (without Velcro) balls for about 1 minute. She allowed the infants to play with the balls and encouraged them to pick the balls up so that the infants could see that the balls were discrete objects.

Demonstration Phase 1. The experimenter brought out the long Plexiglas container with three red balls and three green balls and picked up each ball, one colour at a time, while saying, “Look at this one! See this?” This phase lasted for approximately 30 seconds, after which experimenter cleared the container from the stage.

Familiarization trials. Four familiarization trials were conducted in exactly the same manner as in Experiment 1. These trials lasted a total of approximately 4 minutes.

Demonstration Phase 2. This phase was identical to Experiment 1 except that the balls with Velcro were movable inside the box. The experimenter simply lifted one ball of each colour while saying, “See this one? Look at this!” She completed the sequence twice. This phase lasted approximately 45 seconds.

Test trials. Six test trials were conducted in exactly the same manner as in Experiment 1.

Design

The Design was identical to Experiment 1.

Predictions

We predicted that infants should look longer at a sample that is less probable than one that is more probable given the population of Ping-pong balls in the box. For example, for the 8 infants who saw the mostly red box on test trials, a sample of five green balls is less probable than a sample of five red balls; therefore they should look longer at the five green ball sample. For the 8 infants who saw the mostly green box on test trials, their intuitions should proceed in the opposite direction.[1]

---

[1] The probability of obtaining a sample of 5 green balls from the mostly green box = 60/72 * 59/71 * 58/70 * 57/69 * 56/68 = 0.39034 when sampling without replacement. The probability of obtaining a sample of 5 red balls from the mostly green box = 12/72 * 11/71 * 10/70 * 9/69 * 8/68 = 0.00006 when sampling without replacement. The more probable outcome is 6896 times more likely than the less probable outcome.

Alternatively, if infants do in fact simply look longer at samples of balls with Velcro, they should look longer at the samples with Velcro regardless of the probability of obtaining those balls given the proportions in the population. For half of the infants, the balls with Velcro constituted the majority of balls in the population and for the other half of the infants, the balls with Velcro constituted the minority of balls in the population. Therefore, if infants simply look longer at samples with Velcro, this should result in infants looking about equally at the more probable versus the less probable samples.

Results

Preliminary analyses found no effects of gender, content of the box (mostly green vs. mostly red), colour of balls with Velcro, or order of the familiarization trials. Subsequent analyses collapsed over these variables. A second observer who was unaware of the order of the trials coded the test trials. Inter-observer reliability averaged 96%.

An ANOVA examined the effects of test trial order (expected outcome first or unexpected outcome first), trial pair (1, 2, and 3), and outcome (expected vs. unexpected, *according to the content of the box*). There was a main effect of outcome, $F(1,15) = 7.35$, $p < .05$, effect size (ηp^2) = .33. Infants looked reliably longer at the unexpected outcome (M = 9.81s, SD = .71) than the expected outcome (M = 6.21s, SD = 1.25) (see Figure 3.1). Twelve of sixteen infants looked longer at the unexpected outcome, Wilcoxon sign-ranks test: $z = -2.17$, $p < .05$. There were no other main effects or interactions.

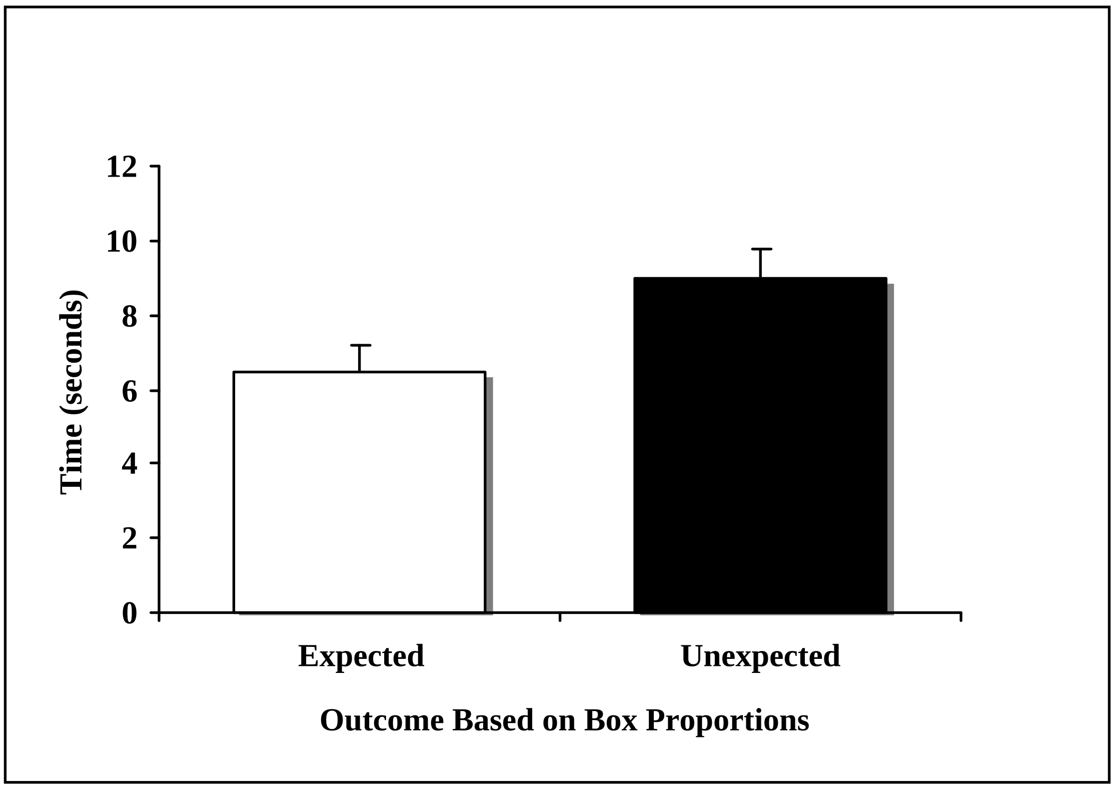
Next we tested the alternative hypothesis that infants simply tend to look longer at samples of balls covered in Velcro. Similarly to Experiment 1, we split infants into two groups. Infants in the Consistent group (n = 8) saw unexpected samples that also happened to consist of balls covered in Velcro (i.e., infants saw a mostly red box with the green balls covered in Velcro). In this group, infants would look longer at a sample of five green balls if they were reasoning about probability or if they were simply inclined to look longer at balls with Velcro. Infants in the Inconsistent group (*n* = 8) saw unexpected samples that consisted of balls not covered in Velcro (i.e., a mostly red box with red balls covered in Velcro). In this group, infants would look longer at a sample of five green balls if reasoning about probability and a sample of five red balls if they were simply looking longer at balls with Velcro. An ANOVA revealed no interaction between consistency and outcome, $F(1,14) = 0.16$, $p > .05$, effect size (ηp^2) = .011. Therefore, infants did not differ in their looking times at the expected and unexpected outcomes *according to the probability of the population* depending on whether or not the balls in the samples were covered in Velcro. This finding makes the interpretation that infants simply looked longer at samples containing balls with Velcro in Experiment 1 extremely unlikely.

Discussion

When no balls were demonstrated to have the property of being stuck, infants looked longer at, for example, a sample of five green balls being removed from a mostly red box than a sample of five red balls being removed from a mostly red box. That is, infants used a probabilistic cue to make inferences about populations given the composition of a sample. This result replicates earlier findings suggesting that 8- and 11-month-old infants can make inferences from samples to populations. The results of this experiment also ruled out the interpretation that infants simply looked longer at the unexpected outcome in Experiment 1 because the physical constraint was correlated with whether or not the balls in the samples were covered in Velcro.

The findings of Experiments 1 and 2 suggest that, when infants are presented with physical knowledge that conflicts with probabilistic knowledge, they can correctly integrate this information. Infants appear to override probabilistic information and instead use information regarding a physical constraint to guide their expectations about samples and populations. Experiment 3 asks whether infants can integrate a physical constraint into this statistical inference task in a more challenging situation. In Experiment 3, we will show infants a large population containing three sets of balls with a ratio of 5 stuck green balls to 4 red balls to 1 yellow ball. We asked whether infants can keep in mind that the green balls are stuck and then estimate which of two samples is more likely—four red balls and one yellow ball or four yellow balls and one red ball. If infants' looking behavior suggests that they find the sample of four yellow balls and one red ball less likely than the sample of four red balls and one yellow ball, this will provide further evidence that infants can integrate domain-specific knowledge into this probabilistic inference mechanism.

Figure 3.1. Average Looking Time at Expected and Unexpected Outcomes in Experiment 2.



4. Experiment 3

Method

Participants

Participants were forty 11-month-old infants (20 girls and 20 boys; mean age = 11;3 [months;days], range = 10;16 to 11;15). Twenty infants (10 girls and 10 boys) were randomly assigned to each of two conditions: the Physical Constraint condition (mean age = 11;01), and the No Physical Constraint condition (mean age = 11;02). An additional six and three infants in each condition were tested but not included in the final sample due to experimenter error (1), parental interference (3), or not completing the study due to fussiness (5). All participants were recruited from the Greater Vancouver area and parents filled out a consent form at the beginning the experiment. Infants received a T-shirt or bib and a diploma for their participation.

Materials and Apparatus

Ping-Pong Balls. A total of 292 (146 green, 73 red and 73 yellow) Ping-pong balls were used. All of the green balls had six white Velcro strips glued to them.

Boxes and Containers. Boxes and containers were the same as those in Experiments 1 and 2 except that there were three colours of Ping-pong balls. In the Physical Constraint condition, the green balls were glued to the boxes and containers and in the No Physical Constraint condition they were not. The small white box used in the Free Play Period and the Plexiglas container used in Demonstration Phase 1 contained two green, two red, and two yellow balls. The box used in Demonstration Phase 2 contained six green, six red, and six yellow balls.

The large box used in the Familiarization Trials and Test Trials had a total of 200 balls with a ratio of 5 green (with Velcro): 4 red: 1 yellow balls on one side and 5 green (with Velcro): 4 yellow: 1 red balls on the other side.

The Apparatus was the same as Experiment 1.

Procedure: Physical Constraint Condition

The procedure of the current experiment was the same as in Experiment 1 with any differences resulting from the fact that there were three colours of balls.

Free Play Period. At the beginning of the experiment, the experimenter showed infants the small white box with two green, two red and two yellow balls for about 1 minute. She allowed the infants to play with the balls and encouraged them to pick the balls up so that the infants could see that the balls were discrete objects and that the green balls were stuck to the box.

Demonstration Phase 1. The experimenter brought out the long Plexiglas container with two green, two red and two yellow balls in random order. She picked up (in counterbalanced order) the red balls one at a time while saying, “Look at this one! See this?” She then did the same with the yellow balls. Then, she grasped the green balls one at a time and said, “Look at this one. It’s stuck!” This phase lasted for approximately 30 seconds, after which experimenter cleared the container from the stage.

Familiarization trials. Four familiarization trials were conducted in exactly the same manner as in Experiment 1. These trials lasted a total of approximately 4 minutes.

Demonstration Phase 2. This phase was identical to Experiment 1 except that the experimenter picked up 1 red ball and 1 yellow ball and grasped one green ball. This phase lasted approximately 45 seconds.

Test trials. Six test trials were conducted in exactly the same manner as in Experiment 1. Infants either saw the box with a 5:4:1 ratio of green to red to yellow box or a 5:4:1 green to yellow to red ratio for all six test trials. On alternating trials, the experimenter pulled four red balls and one yellow ball from the closed box or four yellow balls and one red ball from the closed box. She then lifted the curtain to reveal the population. This phase lasted approximately 7 minutes.

Procedure: No Physical Constraint Condition

The No Physical Constraint condition was exactly the same as the Physical Constraint condition except that all of the balls were movable.

Design: Physical Constraint Condition

Ten infants saw the 5 green: 4 red: 1 yellow balls ratio side of the box during test trials and ten infants saw the box with a ratio of 5 green: 4 yellow: 1 red balls. The order of the familiarization trials (5 green: 4 red: 1 yellow vs. 5 green: 4 yellow: 1 red) and the order of the samples on the test trials (four red and one yellow or four yellow and one red first) were counterbalanced across infants. The order during the Demonstration Phases (green, red or yellow first) was randomized across infants.

Design: No Physical Constraint Condition

The design was the same as the Physical Constraint Condition.

Predictions: Physical Constraint Condition

We predicted that infants would detect the physical constraint (i.e., green balls are immovable) and therefore they should look longer at a sample that is less probable than one that is more probable given the population of Ping-pong balls in the box. For example, for the 10 infants who saw the 5 green: 4 red: 1 yellow ratio of balls in the population on test trials, a sample of four yellow and one red balls is less probable than a sample of four red and one yellow balls.[1] For the 10 infants who saw the 5 green: 4 yellow: 1 red ratio of balls in the population, the opposite is true.

Predictions: No Physical Constraint Condition

Because the green balls were movable, the probability of obtaining either sample (four red and one yellow or four yellow and one red) from a box containing 50% green balls is extremely low. However, it is technically more likely to obtain a sample of four red and one yellow balls than a sample of four yellow and one red balls from a box with a ratio of 5 green: 4 red: 1 yellow balls.[2] Due to the fact that these probabilities are both so low (< 2%), we predicted that infants will find both outcomes unexpected and will look about equally, and possibly for a relatively long time, at each outcome.

---

[1] Infants could carry out either of the following computations in order to succeed at this task:
1) The probability of obtaining a sample of a particular order of 4 yellow and 1 red balls from a box containing 40 yellow and 10 red balls = 40/50 * 39/49 * 38/48 * 37/47 * 10/46 = 0.08627. The probability of obtaining a sample of a particular order of 4 red and 1 yellow balls from this box = 10/50 * 9/49 * 8/48 * 7/47 * 40/46 = 0.00079.
2) The probability of obtaining a sample of 4 yellow and 1 red balls, irrespective of order = 0.08627 *5 = 0.43134. The probability of obtaining a sample of 4 red and 1 yellow balls from this box, irrespective of order = 0.00079 * 5 = 0.00396.
The second set of computations seems a more plausible strategy than the first, although it is possible to succeed at this task via some alternative strategy.
The more probable outcome is 109 times more likely than the less probable outcome for both sets of computations.

[2] 1) The probability of obtaining a sample of a particular order of 4 yellow and 1 red balls from a box with 50 green, 40 yellow and 10 red balls = 40/100 * 39/99 * 38/98 * 37/97 * 10/96 = 0.00243. The probability of obtaining a sample of a particular order of 4 red and 1 yellow balls from this box = 10/100 * 9/99 * 8/98 * 7/97 * 40/96 = 0.00002.
2) The probability of obtaining a sample of 4 yellow and 1 red balls, irrespective of order = 0.0024 * 5 = 0.01214. The probability of obtaining a sample of 4 red and 1 yellow balls from this box, irrespective of order = 0.000022 * 5 = 0.00011.
The more probable outcome is 109 times more likely than the less probable outcome for both sets of computations. However, all four probabilities have less than 2% chance of occurring.

Coding

For the 10 infants who saw the population with a ratio of 5 green: 4 red: 1 yellow during test trials, we coded the sample of four red and one yellow balls as expected and the sample of four yellow and one red balls as unexpected. For the 10 infants who saw the population with a ratio of 5 green: 4 yellow: 1 red during test trials, trials were coded in the opposite way. This coding was used for both conditions, even though we predicted that infants in the No Physical Constraint Condition might find both of the sample types unlikely given that the green balls were not stuck inside the box.

Results

Preliminary analyses found no effects of gender, content of the box on test trials (5 green: 4 red: 1 yellow vs. 5 green: 4 yellow: 1 red), order of samples on test trials (4 red and 1 yellow first or 4 yellow and 1 red first) or order of the familiarization trials. Subsequent analyses collapsed over these variables. A second observer who was unaware of the order of the trials coded a random sample of 10 infants from each condition. Inter-observer reliability averaged 97% and 95% for the two conditions, respectively.

An ANOVA found no differences between average looking times on the familiarization trials across conditions, $F(3,111) = 0.047$, $p > .5$; $M_{physical\ constraint} = 8.58$s, $SD_{physical\ constraint} = 5.01$s. $M_{no\ physical\ constraint} = 10.33$s; $SD_{no\ physical\ constraint} = 5.40$s.

An Omnibus ANOVA examined the effects of test trial order (expected outcome first or unexpected outcome first), trial pair (1, 2, and 3), outcome (expected vs. unexpected) and condition (Physical Constraint vs. No Physical Constraint). There was a main effect of trial pair, $F(2,76) = 5.82$, $p < .05$, average looking decreased over time. There was also an interaction between outcome and condition, F (1,38) = 6.32, p < .05, effect size (ηp^2) = .14 (see Figure 4.1). There were no other main effects or interactions.

Given the interaction between Outcome and Condition and our specific *a priori* looking time predictions, ANOVAs were performed to analyze each condition separately.

Results: Physical Constraint Condition

An ANOVA examined the effects of test trial order (expected outcome first or unexpected outcome first), trial pair (1, 2, and 3), and outcome (expected vs. unexpected). There was a main effect of outcome, F(1,19) = 14.66, p < .05, effect size (ηp^2) = .44. Infants looked reliably longer at the unexpected outcome (*M* = 7.45s, *SD* = 3.70s) than the expected outcome (*M* = 5.53s, *SD* = 3.85s). Sixteen of twenty infants looked longer at the unexpected outcome, Wilcoxon signed-ranks test: z = -2.91, p < .05. There were no other main effects or interactions.

Results: No Physical Constraint Condition

An ANOVA examined the effects of test trial order (expected outcome first or unexpected outcome first), trial pair (1, 2, and 3), and outcome (expected vs. unexpected). There was a main effect of trial pair, F(2,38) = 4.37, p < .05; average looking decreased over time. There was no main effect of outcome, F (1,19) = .21, p > .5, effect size (ηp^2) = .01. Infants looked about equally at the unexpected outcome (*M* = 8.07s, *SD* = 4.03) and the expected outcome (*M* = 7.72, *SD* = 4.00). Nine of twenty infants looked longer at the unexpected outcome, Wilcoxon signed-ranks test: z = -.23, p > .5. There were no other main effects or interactions.

To test the hypothesis that both outcomes in the No Physical Constraint Condition were unexpected, we conducted four post-hoc t-tests comparing the mean looking times obtained from the Physical Constraint and No Physical Constraint Conditions: First, we compared the expected outcome in the Physical Constraint condition ($M = 5.53$s) with the "expected" outcome in the No Physical Constraint condition ($M = 7.72$s), $t(38) = -2.10$, $p = .02$. Second, we compared the expected outcome in the Physical Constraint condition ($M = 5.53$s) with the "unexpected" outcome in the No Physical Constraint condition, ($M = 8.07$s), $t(38) = -2.12$, $p = .02$.[3] Third, we compared the unexpected outcome in the Physical Constraint condition ($M = 7.45$s) with the "expected" outcome in the No Physical Constraint condition ($M = 7.72$s), $t(38) = -0.50$, $p > .5$. Finally, we compared the unexpected outcome in the Physical Constraint condition ($M = 7.45$s) with the "unexpected" outcome in the No Physical Constraint condition ($M = 8.07$s), $t(38) = -0.24$, $p > .5$. These results indicate that the mean looking times for the unexpected outcome in the Physical Constraint condition and the "expected" and "unexpected" outcomes in the No Physical Constraint condition were not reliably different from one another, suggesting that both outcomes in the No Physical constraint condition were unexpected. The expected outcome for the Physical Constraint condition was different from each of the other three means.

Discussion

In the Physical Constraint condition, infants were given evidence that green balls were physically constrained inside the box. They looked longer at a sample of, for example, four yellow balls and one red ball than a sample of four red balls and one yellow ball from a box with a ratio of 5 stuck green balls to 4 red balls to 1 white ball. This suggests that infants kept in mind that green balls were physically constrained and could thus be disregarded from probability calculations. Excluding the green balls allowed infants to correctly make

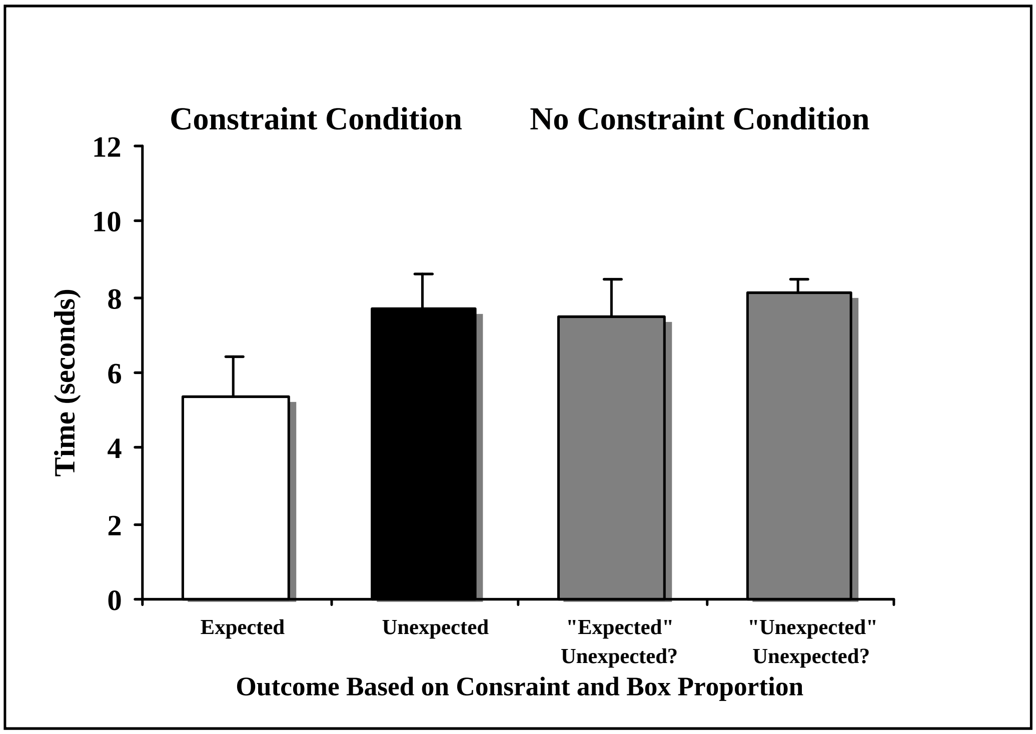
[3] According to the Bonferroni technique, these *p* values are marginally significant (Bonferroni requires an alpha level of .0125 for these tests).

inferences regarding the likelihood of a mostly yellow or mostly red sample from a box with either a greater proportion of red or yellow balls.

In contrast, infants in the No Physical Constraint condition were not given evidence that green balls were stuck inside the box. They looked about equally at both samples. This suggests that infants did not find either the unexpected outcome or the expected outcome less probable in this condition.

Furthermore, the analyses indicating that there were no reliable differences in looking times at both outcomes in the No Physical Constraint condition and the unexpected outcome in the Physical Constraint condition suggest that infants may have found all three of these outcomes improbable. It is also possible, however, that infants in the No Physical Constraint condition were simply unable to reason about the probability of three sets of balls without the benefit of excluding one of the sets. This may have caused them to look equally at the two samples in this condition. Teasing apart whether infants found both of these outcomes unexpected or they were simply unable to reason about three sets of balls is an empirical question regarding the computational limits of this mechanism. However, the fact that infants' looking times on these trials were relatively long and were similar in length to the looking time on the unexpected trials in the Physical Constraint condition makes this alternative interpretation less plausible. Regardless, findings from the No Physical Constraint condition control for the possibility that infants in the Physical Constraint condition did not detect the physical constraint and were instead computing which of two very unlikely probabilities was least likely. Taken together, the results of these two conditions suggest that infants in the Physical Constraint condition were sensitive to, and integrated, a physical constraint into the statistical inference mechanism.

Figure 4.1. Average Looking Time at Expected and Unexpected Outcomes in Experiment 3.



5. General Discussion

We conducted three experiments to investigate whether or not infants can integrate a physical constraint into a statistical inference mechanism. In Experiment 1, we demonstrated to infants that one set of balls, those covered in Velcro, was stuck inside a box. We then removed a sample of either five balls with Velcro or five balls without Velcro from a closed box. We found that infants looked longer at a sample containing five balls with Velcro, regardless of whether the probability of obtaining balls with Velcro was high or low based on the proportions in the box. In Experiment 2, we replicated earlier findings suggesting that infants can make inferences from samples to populations. We showed infants samples of either five red or five green balls being drawn from a box and timed how long they looked at, for example, a population of 5 green to 1 red balls. Infants looked longer at the less probable outcome—a sample of five red balls from the mostly green box. These results also rule out the alternative interpretation that infants looked longer at the unexpected samples in Experiment 1 due to the fact that balls with Velcro were more interesting to look at than balls without Velcro. Finally, in Experiment 3, we demonstrated that balls with Velcro had the property of being physically constrained for half of the infants and that balls with Velcro moved freely for the other half. Then infants saw a population of, for example, 5 green (with Velcro) to 4 red to 1 yellow balls. When infants were shown that green balls with Velcro were stuck, they looked longer at a sample of four yellow balls and one red ball being drawn from the box than a sample of four red balls and one yellow ball. In contrast, when infants were not shown that green balls with Velcro were stuck, they looked about equally at both of these samples.

The remainder of this discussion consists of four sections. In the first section, we will discuss our rationale for proposing that the physical constraint violated in this task is one of cohesion. The second section considers how the results of these experiments fit with existing evidence regarding the integration of domain-specific knowledge and probabilistic inference mechanisms. In the third section, we will consider the implications of these experiments on the nature of the mechanism involved in this statistical inference task. Finally, we will conclude by proposing a number of directions for future research.

The Principle of Cohesion

Research on infants' understanding of physical knowledge suggests that they have a robust grasp of the principles of cohesion, continuity, and contact by the end of the first year of life (Ball, 1973; Leslie, 1988, 1994; Needham, 1999; Spelke, 1990, 1994; Spelke & Born, 1983; Spelke et al, 1992,1995; Spelke & Van de Walle, 1995). The current experiments focused on the principle of cohesion, a constraint that infants are sensitive to at as young as 4 months of age. Given that a violation of cohesion can occur in two ways—objects should constantly maintain both their boundaries and their connectedness, our physical constraint actually consists of two violations of cohesion.

First, we violated infants' understanding of boundedness, which suggests that discrete objects such as balls and boxes should not merge together (Spelke, 1994). We gave infants evidence that a set of balls was stuck inside a box. The property of balls being stuck to boxes is a violation of cohesion, given that under normal circumstances, infants and adults would conceptualize a collection of balls as discrete objects (Needham & Baillargeon, 1997). We demonstrated the property to infants in a variety of ways, showing them that they themselves could not lift the balls from the bottom of a box and that the experimenter could not lift the balls from a box or a container. We also showed them that the balls would not move inside a box even when it was shaken vigorously. After viewing all of these demonstrations, infants presumably were convinced that the balls were no longer discrete objects but instead they were bounded to the insides of boxes.

In Experiment 1, after providing the above evidence regarding a physical constraint, we violated cohesion in a second way, by violating infants' understanding of connectedness— that objects do not spontaneously break apart. We either removed a sample of balls from the box that were not demonstrated to be stuck inside the box or we removed a sample of balls from the box that were demonstrated to be stuck inside the box. Infants looked longer at the sample of balls that were shown to be stuck, suggesting that they detected this violation of connectedness. The fact that infants used their knowledge of cohesion to override probabilistic knowledge suggests that they correctly integrated these two sources of information in this task. In Experiment 3, we did not violate connectedness (i.e., no green balls exited the box) and instead infants were required to integrate the physical constraint into the probabilistic inference mechanism by excluding the set of balls that were constrained. The fact that infants could integrate this knowledge suggests that they are able to reason about the behaviour of physical objects and probability simultaneously. If infants thought that balls could simply break apart from the box (i.e., violate connectedness), or they could not detect the physical constraint demonstrated to begin with (i.e., a violation of boundedness), there would have been no differences in the looking-time patterns of infants in the Physical Constraint and the No Physical Constraint conditions of Experiment 3.

Integration of Domain-Specific Knowledge and Statistical Inference

According to findings from the current experiments and findings from Teglas, Girotto, Gonzalez, & Bonatti, L. (2007), towards the end of the first year, infants can integrate two principles of object knowledge into probabilistic inference mechanisms. Teglas et al. found that infants overrode probabilistic information in favor of their knowledge regarding solidity, which states that objects cannot pass through other solid objects. In their experiment, infants looked longer when a more probable object versus a less probable object (according to the proportions in the population) exited a lottery machine only if the less probable object was prohibited from exiting the machine by a solid barrier. This suggests that infants correctly integrated the probabilistic and physical evidence in the display by reasoning in accord with a violation of solidity. This fits nicely with Experiment 1 of the current experiments, which suggests that infants will override probabilistic information with physical object knowledge when faced with a violation of cohesion.

The experiment conducted by Xu and Denison (in press) investigated infants' ability to integrate a psychological constraint into the statistical inference mechanism used in the current experiments. A great deal of evidence suggests that infants can reason about the goal-directed actions of human agents (Biro & Leslie, 2007; Hofer, Hauf, & Aschersleben, 2005; Luo & Baillargeon, 2005; Woodward, 1998). Xu and Denison (in press) explored whether or not 11-month-old infants were sensitive to factors such as an agent's goal and whether or not the agent had visual access to a population while sampling. The results of that experiment suggest that infants can integrate probabilistic and psychological knowledge under appropriate sampling conditions. That is, if an experimenter has visual access to a population of balls during sampling, the sample removed from the box should accord with the experimenter’s preference and not with what would constitute a random sample. Therefore, similar to Experiment 1 of the current experiments, infants overrode the probabilistic information in the population and reasoned in accord with the psychological constraint (i.e., the experimenter’s goal).

The results of Teglas et. al. (2007), Xu and Denison (in press), and the current experiments suggest that infants can correctly integrate both physical and psychological knowledge into statistical inference mechanisms. However, the tasks used in Teglas et al. and Xu and Denison did not require infants to integrate their domain-specific knowledge into the statistical inference mechanisms and then compute probabilities over a remaining subset of objects. Infants could succeed at these experiments, as well as in Experiment 1 of the current experiments without computing any probabilities; they only had to exclude the correct set of objects. Experiment 3 adds to the evidence suggesting that infants can integrate domain-specific knowledge into a domain-general inference mechanism by requiring infants to exclude a set of objects via physical reasoning and then compute probabilities over two sets of remaining objects.

The Nature of the Probabilistic Mechanism

The current experiments also shed some light on the characteristics of this statistical inference mechanism. The explored mechanism appears to be available to infants as young as 8-months of age, and possibly younger. Currently, evidence exists suggesting that 6-month-olds are not yet capable of using this mechanism (Xu, Garcia, & Waechtler, unpublished data). Infants were shown displays similar to the ones used with 8-month-olds (Xu & Garcia, 2008); they watched an experimenter draw a sample of four red balls and one white ball or a sample of four white balls and one red ball from a box containing mostly red balls. Six-month-old infants did look reliably longer at the less probable outcome. However, in a control condition, they also looked longer at a sample of four white balls and one red ball placed next to a mostly red box when the balls were drawn from the experimenter’s pocket rather than from the box. This suggests that 6-month-old infants may not be making inferences from samples to populations but instead they may simply look longer at mismatches between the two. Therefore, this mechanism appears to become available to infants some time between 6 and 8 months of age.

Due to the fact that 8-month-old infants looked longer at the unexpected outcome in the experimental condition but they looked about equally at both outcomes in the control condition, it seems unlikely that older infants are simply making perceptual matches between samples and populations without considering whether the sample came from the population or not. However, Xu and Garcia (2008) left open the question of whether infants are computing probabilities over the input or using a representativeness heuristic (namely, that samples and populations should resemble one another in appearance) to make inferences. Results from the Physical Constraint condition of Experiment 3 suggest that, at least in one context, infants are probably not solely relying on representativeness to reason about the relationship between samples and populations. Here the sample of four red balls and one yellow ball and the sample of four yellow balls and one red ball did not resemble the population perceptually, nonetheless infants looked longer at the less probable sample. Therefore, it seems plausible that infants may be computing or estimating the probability of obtaining the samples based on the composition of the population and not solely reasoning that a sample drawn from a particular population should match it perceptually. The interpretation that infants may be excluding the green balls and selectively attending to the red and yellow balls in the box is a less likely explanation given that the colours of balls in the box are intermixed and not separated into distinct regions (although, see Halberda, Sires, & Feigenson (2006) for evidence that adults will selectively attend to large sets of intermixed coloured dots).

Future Directions

The current experiments leave open a number of questions for future research. First, we now have evidence that 11-month-old infants can integrate their understanding of cohesion and agents, and 12-month-old infants can integrate solidity, into statistical inference mechanisms under some conditions. However, more evidence is required to strengthen the claim that infants are capable of integrating domain-specific knowledge into this task. Can younger infants integrate knowledge similarly to older infants? Are there other physical and psychological constraints that can be integrated into the mechanism?

The first open question concerns the developmental trend of infants' abilities to integrate domain-specific knowledge and the statistical mechanism used in the current experiments. We currently have evidence that this mechanism is available to 8-month-old infants, but we do not know whether or not they can combine this task with other sources of knowledge. It seems plausible that these younger infants who possess a fairly robust understanding of the principle of cohesion might perform well in the current experiments. However, it also seems plausible that they might not succeed at integrating the psychological constraint utilized in Xu and Denison (in press). Although there is evidence that infants as young as 5-months of age understand simple goal directed actions, they may not appreciate visual access as a cue to whether or not an agent can perform a goal-directed action (Woodward, 1998). Therefore, 8-month-old infants should be able to integrate physical constraints of which they have a firm grasp into this inference mechanism but might not be able to integrate more advanced knowledge of an agent's behaviour into the task. If 8-month-old infants cannot integrate a physical constraint into this mechanism, this might suggest that their ability to use the mechanism is too fragile, or that the cognitive demand of keeping in mind a physical constraint and excluding a set of objects from a probability calculation is too high.

Another question addresses whether or not other physical constraints can be integrated into this mechanism. Because Teglas and colleagues found that 12-month-old infants can integrate a violation of solidity into a probabilistic task, we propose a future experiment that will test whether infants can integrate a violation of solidity into the current statistical inference mechanism. Previous experiments investigated infants' understanding of solidity by habituating infants to a scene where a ball was dropped behind an occluder. On test trials, infants dishabituated to an event where the ball appeared on the table below a platform containing a hole too narrow for the ball to fit through (Spelke, 1994). In a future experiment, we will display a population of balls of different sizes in a box with a hole that is large enough to fit some of the balls but not large enough to fit others. We will then show infants test events that directly set probabilistic information in opposition with their physical knowledge of solidity by drawing two different samples from the box on alternating test trials. One sample will be more probable based on the proportion of balls in the box but impossible to obtain based on the size of the hole in the box and the size of the balls. The other sample will be less probable based on the proportion of balls in the box but possible based on the size of the hole and the size of the balls. If infants are capable of integrating solidity into this probabilistic inference mechanism, results will provide converging evidence regarding integration of solidity in this method and Teglas et. al's method. This will also produce a second piece of evidence suggesting that infants can integrate physical knowledge into this statistical inference mechanism.

Finally, a number of future experiments are required to establish the computational limits of this mechanism. In the No Physical Constraint condition in Experiment 3, infants looked about equally at two different samples even though one of the two samples was actually more probable than the other. We proposed that infants did not differentiate between the two outcomes because they were both extremely low in probability. These results raise the question of whether or not infants are capable of detecting differences in low probability outcomes in general.

There are also at least two alternative explanations for the findings in the No Physical Constraint condition in Experiment 3. First, it is possible that our measures are not sensitive enough to detect whether or not infants found one of these two outcomes less probable than the other. We believe that this interpretation is relatively unlikely given that the differences in the probability of obtaining the more probable sample in the Physical Constraint condition and the No Physical constraint condition were exactly the same (i.e., roughly 109 times more likely). Second, infants may simply be unable to reason about probability when faced with three sets of balls, and thus they looked equally at the two outcomes. We believe that this explanation is less plausible than the account suggesting that infants found both outcomes unlikely, given that infants looked for a relatively long duration at these two outcomes. The long duration of looking demonstrated by infants provides tentative evidence that they may have found both outcomes improbable. Nevertheless, this raises the question of whether or not infants are capable of computing probabilities over three sets of balls. For example, if infants were shown a box with a ratio of 5 green to 4 red to 1 yellow ball and then saw samples of two green balls, two red balls and one yellow ball versus four red balls and one yellow ball being drawn from the box, would they look longer at the less probable event? Research suggests that the adult visual system can keep track of a maximum of three sets of objects in large arrays (Alvarez & Cavanagh, 2004; Halberda et al, 2006; Luck & Vogel, 1998; vanMarle & Scholl, 2003; Wynn, Bloom, & Chiang, 2002). This suggests that three sets of balls may be an upper limit that infants cannot exceed as well; however, it is also possible that infants can only make inferences from samples to populations when just two sets of balls are present. If three sets of balls is in fact an upper limit, could infants and adults overcome this limit if they were presented with four sets of balls but they were able to exclude one set from probability calculations based on a physical or psychological constraint? These are empirical question that can only be answered via further exploration of this inferential mechanism.

Conclusion

The current experiments explored a statistical inference task in 11-month-old infants. Results replicated earlier findings, showing that infants are capable of making inferences from samples to populations. Our findings also suggest that, when substantive domain-specific knowledge is in conflict with probabilistic information, infants override probabilistic information and reason according to the constraint. Furthermore, we found that infants can use a physical constraint to exclude one set of balls from probability calculations and then compute probabilities over two remaining sets of balls. These results, combined with findings suggesting that infants can integrate psychological knowledge into this statistical inference task, provide further evidence that infants can meaningfully combine these two sources of information. In general, the findings of this and related experiments demonstrate that the study of learning mechanisms and domain-specific knowledge should go hand in hand.

References

Aguiar, A. & Baillargeon, R. (2002). Developments in young infants' reasoning about occluded objects. *Cognitive Psychology, 45*, 267-336.

Alvarez, G. A., & Cavanagh, P. (2004). The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychological Science, 15(2),* 106–111.

Aslin, R., Saffran, J. & Newport, E. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological Science, 9,* 321-324.

Baillargeon, R. (2008). Innate ideas revisited: For a principle of persistence in infants' physical reasoning. *Perspectives on Psychological Science, 3(1),* 2-13.

Baillargeon, R. (2004) Infants' physical world. *Current Directions in Psychological Science, 3,* 89-94.

Ball, W.A. (1973). The perception of causality in the infant. *Paper presented at the Society for Research in Child Development.* Philadelphia.

Biro, S., & Leslie, A.M. (2007). Infants' perception of goal-directed actions: Development through cue-based bootstrapping. *Developmental Science, 10*, 379–398.

Carey, S. (in press). *The origins of concepts.* Oxford University Press.

Carey, S., & Spelke, E. (1994). Domain-specific knowledge and conceptual change. *Mapping the mind: Domain specificity in cognition and culture* (pp. 169-200). New York, NY US: Cambridge University Press.

Cosmides, L., & Tooby, J. (1994). Origins of domain specificity: The evolution of functional organization. *Mapping the mind: Domain specificity in cognition and culture* (pp. 85-116). New York, NY US: Cambridge University Press.

Elman, J., Bates, E., Johnson, M., & Karmiloff-Smith, A. (1996). *Rethinking innateness: A connectionist perspective on development.* Cambridge, MA US: The MIT Press.

Gergely, G., Nádasdy, Z., Csibra, G., & Bíró, S. (1995). Taking the intentional stance at 12 months of age. *Cognition*, *56,* 165-193.

Gerken, L. (2006). Decisions, decisions: infant language learning when multiple generalizations are possible. *Cognition, 98,* B67-B74.

Gopnik, A., Glymour, C., Sobel, D., Schulz, L., Kushnir, T. & Danks, D. (2004). A theory of causal learning in children: causal maps and Bayes nets. *Psychological Review, 111,* 1-31.

Gopnik, A., Sobel, D., Schulz, L., & Glymour, C. (2001). Causal learning mechanisms in very young children: Two-, three-, and four-year-olds infer causal relations from patterns of variation and covariation. *Developmental Psychology, 37(5),* 620-629.

Gomez, R.L. (2002). Variability and detection of invariant structure. *Psychological Science, 13,* 431-436.

Halberda, J., Sires, S., & Feigenson, L. (2006). Multiple Spatially Overlapping Sets Can Be Enumerated in Parallel. *Psychological Science, 17(7),* 572-576.

Hespos, S., & Baillargeon, R. (2001). Reasoning about containment events in very young infants. *Cognition, 78(3),* 207-245.

Hirschfeld, L., & Gelman, S. (1994). *Mapping the mind: Domain specificity in cognition and culture.* New York, NY US: Cambridge University Press.

Hofer, T., Hauf, P., & Aschersleben, G. (2005). Infant's perception of goal-directed actions performed by a mechanical device. *Infant Behavior & Development, 28,* 466–480.

Karmiloff-Smith, A. (1992). Nature, nurture and PDP: Preposterous Developmental Postulates? *Connection Science, 4(3)*, 253-269.

Keil, F. (1989). *Concepts, kinds, and cognitive development*. Cambridge, MA US: The MIT Press.

Kushnir, T. & Gopnik, A. (2007). Conditional probability versus spatial contiguity in causal learning: Preschoolers use new contingency evidence to overcome prior spatial assumptions. *Developmental Psychology, 43(1)*, 186-196.

Leslie. A.M. (1994). ToMM, ToBy, and agency: Core architecture and domain specificity. In L.

Hirschfeld & S. Gelman (Eds.), *Domain-specificity in cognition and culture.* New York: Cambridge University Press.

Leslie. A.M. (1988). The necessity of illusion: Perception and thought in infancy. In L. Weiskrantz (Ed.). *Thought without language.* Oxford: Clarendon Press.

Leslie, A.M., & Keeble, S. (1987). Do six-month-old infants perceive causality? *Cognition, 25,* 265–288.

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

Luo, Y., & Baillargeon, R. (2005). Can a self-propelled box have a goal? Psychological reasoning in 5-month-old infants. *Psychological Science, 16*, 601–608.

Maye, J., Werker, J., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition, 82(3),* B101-B111.

Marcus, G. F., Vijayan, S., Bandi Rao, S., & Vishton, P.M. (1999). Rule learning by seven-month-old infants. *Science, 283,* 77-80.

Needham, A. (1999). The role of shape in 4-month-old infants' object segregation. *Infant Behavior & Development*, *22(2),* 161-178.

Needham, A. & Baillargeon, R. (1997). Object segregation in 8-month-old infants. *Cognition, 62*, 121-129.

Newport, E.L., & Aslin, R.N. (2004). Learning at a distance: I. Statistical learning of non-adjacent dependencies. *Cognitive Psychology, 48*, 127-162.

Pinker, S. (1994). *The language instinct.* New York, NY US: William Morrow & Co.

Pinto, J. (2005). Xhab1.4. Stanford University.

Quinn, P. (1987). The categorical representation of visual pattern information by young infants. *Cognition, 27(2),* 145-179.

Saffran, J., Aslin, R. & Newport, E. (1996). Statistical learning by 8-month-old infants. *Science, 274,* 1926-2928.

Saffran, J., Newport, E., Aslin, R., & Tunick, R. (1997). Incidental language learning: Listening (and learning) out of the corner of your ear. *Psychological Science, 8(2),* 101-105.

Schmidt, H., Spelke, E., & LaMorte, V. (1986). The development of Gestalt perception in infancy. *Paper presented at the International Conference on Infant Studies*, Los Angeles.

Schulz, L.E., Bonawitz, E. B., & Griffiths, T. (2007). Can being scared make your tummyache? Naive theories, ambiguous evidence and preschoolers' causal inferences. *Developmental Psychology 43*(5), 1124-1139.

Schulz, L.E. & Gopnik, A. (2004). Causal learning across domains. *Developmental Psychology, 40,* 162-176.

Smith, L. B., Jones, S. S., Landau, B., Gershkoff-Stowe, L., & Samuelson, L. (2002). Object name learning provides on-the-job training for attention. *Psychological Science, 13,* 13-19.

Sobel, D., & Kirkham, N. (2007). Bayes nets and babies: Infants' developing statistical reasoning abilities and their representation of causal knowledge. *Developmental Science, 10(3),* 298-306.

Sobel, D. M. & Kirkham, N. Z. (2006). Blickets and babies: The development of causal reasoning in toddlers and infants. *Developmental Psychology, 42,* 1103-1115.

Spelke, E.S. (1994). Initial knowledge: Six suggestions. *Cognition, 50(1),* 431-445.

Spelke, E.S. (1990). Principles of object perception. *Cognitive Science, 14*, 29-56.

Spelke, E.S. & Born, W.S. (1983). Visual perception of objects by three-month-old infants. Unpublished Manuscript.

Spelke, E., Breinlinger, K., Macomber, J., & Jacobson, K. (1992). Origins of knowledge. *Psychological Review, 99(4),* 605-632.

Spelke, E., Kestenbaum, R., Simons, D., & Wein, D. (1995). Spatiotemporal continuity, smoothness of motion and object identity in infancy. *British Journal of Developmental Psychology, 13(2),* 113-142.

Spelke. E.S., & Van de Walle, G. (1995). Perceiving and reasoning about objects: Insights from infants. In N. Eilan, W. Brewer, & R. McCarthy (Eds.). *Spatial representations.* New York: Basil Blackwell.

Teglas, E., Girotto, V., Gonzalez, M., Bonatti, L. (2007). Intuitions of probabilities shape expectations about the future at 12 months and beyond. *Proceedings of the National Academy of Sciences of the United States of America, 104,* 19156-19159.

VanMarle, K., & Scholl, B. J. (2003). Attentive tracking of objects versus substances. *Psychological Science*, *14(5)*, 498-504.

Wang, S., Baillargeon, R., & Paterson, S. (2005). Detecting continuity violations in infancy: A new account and new evidence from covering and tube events. *Cognition, 95(2),* 129-173.

Woodward, A.L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition, 69*, 1–34.

Wynn, K., Bloom, P., & Chiang, W.-C. (2002). Enumeration of collections by 5-month-old infants. *Cognition*, *83*, B55-B62.

Xu, F. & Denison, S. (in press). Statistical inference and sensitivity to sampling in 11-month-old infants. *Cognition*.

Xu, F. & Garcia, V. (2008). Intuitive statistics by 8-month-old infants. *Proceedings of the National Academy of Sciences of the United States of America*, *105,* 5012-5015.

Xu, F. & Spelke, E.S. (2000). Large number discrimination in 6-month-old infants. *Cognition*, *74*, B1-B11.

Younger, B. (1993). Understanding category members as 'the same sort of thing': Explicit categorization in ten-month infants. *Child Development, 64(1)*, 309-320.

Appendix 1: List of Publications

A combined version of chapters two through four will be submitted for publication.

The University of British Columbia
Office of Research Services and Administration
Behavioural Research Ethics Board

Appendix 2

Certificate of Approval

PRINCIPAL INVESTIGATOR	DEPARTMENT	NUMBER
Xu, F.	Psychology	B03-0657

INSTITUTION(S) WHERE RESEARCH WILL BE CARRIED OUT

UBC Campus ,

CO-INVESTIGATORS:

Atwal, Pritpal, Psychology; Boak, Heather, ; Davies, Heather, ; Denison, Stephanie, ; Dewar, Kathryn, Psychology; Huxtable, Angela, ; Konopczynski, Katie, Psychology; Narayan, Jane, Psychology; Omaid, Tara, Psychology; Perfors, Amy, Psychology

SPONSORING AGENCIES

Natural Science Engineering Research Council

TITLE :

Language and Cognition in Infants and Children

APPROVAL DATE	TERM (YEARS)	AMENDMENT:	AMENDMENT APPROVED:
06-03-13 (yr/mo/day)	1	Nov. 1, 2006, Consent form / Oct. 24, 2006, Recruitment method / Reimbursement	NOV 0 9 2006

CERTIFICATION:

The request for continuing review of an amendment to the above-named project has been reviewed and the procedures were found to be acceptable on ethical grounds for research involving human subjects.

Approved on behalf of the Behavioural Research Ethics Board
by one of the following:
Dr. Peter Suedfeld, Chair,
Dr. Jim Rupert, Associate Chair
Dr. Arminee Kazanjian, Associate Chair
Dr. M. Judith Lynam, Associate Chair

This Certificate of Approval is valid for the above term provided there is no change in the experimental procedures